

From: Chan, Christina
Sent: Friday, October 28, 2005 3:07 PM
To: Basi, Nirmal; STIC-Biotech/ChemLib
Subject: RE: Rush search for 09/823,069

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

RECEIVED
OCT 28 2005
ECH/OIEN DIVISION
(STIC)

-----Original Message-----

From: Basi, Nirmal
Sent: Friday, October 28, 2005 3:01 PM
To: Chan, Christina
Subject: Rush search for 09/823,069

Christina I am seeking approval for a RUSH sequence search for an Amended case, as indicated below. If approved, could you please forward the search to STIC and cc a copy to me.

Examiner: Nirmal S. Basi
Art Unit 1646
Office: Remsen Building, Room 4D68
Mail Room: Remsen Building, room 4C70

Sequence search:

App. #: 09/823,069

Result format: Paper.

Title: **METHODS AND COMPOSITIONS UTILIZING AN ALTERNATIVE SPLICE VARIANT OF THE SIGMA-1 RECEPTOR**

206 1-579 NA
2-192 AA

Inventors: Wheeler, Kenneth

Priority Date: 4/3/2000
Please search:

i) SEQ ID NOs: 1 and 2

10/31/05
JTH

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Date completed:	Search Site	Vendors
Searcher: <u>Beverly e 2528</u>	<input type="checkbox"/> STIC	<input type="checkbox"/> IG
Terminal time: _____	<input type="checkbox"/> CM-1	<input type="checkbox"/> STN
Elapsed time: _____	<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
CPU time: _____	Type of Search	<input type="checkbox"/> APS
Total time: _____	<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> Geninfo
Number of Searches: _____	<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> SDC
Number of Databases: _____	<input type="checkbox"/> Structure	<input type="checkbox"/> DARC/Questel
	<input type="checkbox"/> Bibliographic	<input checked="" type="checkbox"/> Other <u>CGN</u>

TO-1560 (9-90)

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 04:28:43 ; Search time 176 Seconds
(without alignments)
558.632 Million cell updates/sec

Title: US-09-823-069A-2
Perfect score: 1016
Sequence: 1 MQWAGRRWAWAALLAVAA.....RSVARGRLLELTTLFGQDP 192

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1016	100.0	192	Q9NYX0	Q9NYX0 homo sapien
2	990.5	97.5	223	Q99720	Q99720 homo sapien
3	940.5	92.6	223	Q9ROC9	Q9ROC9 rattus norv
4	934.5	92.0	223	Q9Z2W2	Q9Z2W2 rattus norv
5	925	91.0	192	Q9JKU9	Q9JKU9 mus musculu
6	918.5	90.4	223	Q60492	Q60492 cavia porce
7	903.5	88.9	223	O55242	O55242 mus musculu
8	767.5	75.5	187	Q8N7H3	Q8N7H3 homo sapien
9	636.5	62.6	221	Q6DCU6	Q6DCU6 xenopus lae
10	631.5	62.2	223	Q645J3	Q645J3 taricha gra
11	629.5	62.0	221	Q661M1	Q661M1 xenopus tro
12	616.5	60.7	222	Q7ZMG9	Q7ZMG9 brachydanio
13	544	53.5	106	Q7Z653	Q7Z653 homo sapien
14	508	50.0	106	Q9RLJ7	Q9RLJ7 rattus norv
15	228.5	22.5	256	Q9C2G3	Q9C2G3 neurospora
16	227	22.3	224	Q6FKL1	Q6FKL1 candida gla
17	226.5	22.3	256	1 ERG2 NEUCR	Q92254 neurospora
18	224	22.0	217	Q6BSC9	Q6BSC9 debaryomyce
19	218	21.5	223	Q6CL22	Q6CL22 kluyveromyce
20	217.5	21.4	218	Q6C3U4	Q6C3U4 yarrowia li
21	214.5	21.1	218	Q6CEA6	Q6CEA6 yarrowia li
22	212.5	20.9	222	1 ERG2 YEAST	P32352 saccharomyc
23	211	20.8	222	Q755O1	Q755O1 ashbya goss
24	210.5	20.7	221	1 ERG2 MAGGR	P33281 magnaporthe
25	206	20.3	219	1 ERG2 SCHPO	P87113 schizosacch
26	204	20.1	285	2 Q7RVD2	Q7RVD2 neurospora
27	183.5	18.1	241	1 ERG2 USTMA	P32360 ustilago ma
28	94	9.3	393	2 Q9ZVZ4	Q9ZVZ4 arabidopsis
29	94	9.3	527	2 Q6GM93	Q6GM93 xenopus lae
30	93	9.2	397	2 Q8PIU1	Q8PIU1 xanthomonas
31	93	9.2	417	2 Q7NSI9	Q7NSI9 chromobacte

ALIGNMENTS

RESULT 1

Q9NYX0 PRELIMINARY; PRT; 192 AA.
AC Q9NYX0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Sigma 1 receptor beta variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RA Wang L.-M., Shelness G.S., Childers S.R., Mach R.H., Wheeler K.T.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF226604; AAF64280.1; -
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:000247; F:C-8 sterol isomerase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006696; P:ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_Sireceptlik.
DR Pfam; PF04622; ERG2_Sigma1R; 1.
KW Receptor.
SQ SEQUENCE 192 AA; 21483 MW; 9E0C6A120A89126F CRC64;

Query Match 100.0%; Score 1016; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.4e-84;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQWAGRRWAWAALLAVAAVLTVVWLGTQSFVQREIEIAQLARQYAGLDHDLAFSR 60
DB 1 MQWAGRRWAWAALLAVAAVLTVVWLGTQSFVQREIEIAQLARQYAGLDHDLAFSR 60
QY 61 LIVELRLPHGHVLPDEELQWVFVNAAGWGMANCLLHASLSEYVLLFGTALGRHSGET 120
DB 61 LIVELRLPHGHVLPDEELQWVFVNAAGWGMANCLLHASLSEYVLLFGTALGRHSGET 120
QY 121 VHGPGEATAVENGPNTWVVEYGRGVIPTSLFALADTVFSTODFLTPVTLRSYARGLR 180
DB 121 VHGPGEATAVENGPNTWVVEYGRGVIPTSLFALADTVFSTODFLTPVTLRSYARGLR 180
QY 181 LEITTVLFGQDP 192
DB 181 LEITTVLFGQDP 192
RESULT 2
Q99720 PRELIMINARY; PRT; 223 AA.
AC Q99720; O00673; O00725;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)

KW	Receptor.
SQ	SEQUENCE 223 AA; 25128 MW; 0C49863BEB1C443 CRC64;
	Query Match 97.5%; Score 990.5; DB 2; Length 223;
	Best Local Similarity 86.1%; Pred. No. 5.9e-82;
	Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1
QY	1 MQAVGRRWAALLLAVAAVLTVVWLTLGTQSFVFQREIEIQAOLARQYAGLDHDLAFSR 60
DB	1 MQAVGRRWAALLLAVAAVLTVVWLTLGTQSFVFQREIEIQAOLARQYAGLDHDLAFSR 60
QY	61 LIVELRLRHGHVLPDEELQVFNACGWCMCLLLHASISEYVLLFGTALGSRGHS--- 117
DB	61 LIVELRLRHGHVLPDEELQVFNACGWCMCLLLHASISEYVLLFGTALGSRGSGRY 120
QY	118 -----GETVVHGPGCEATAVENGPNTMMVEYGRGVIPS 149
DB	121 WAEISDTIIISGFTHQREGTKSEVPFGETVVHGPGCEATAVENGPNTMMVEYGRGVIPS 180
QY	150 TLFALADTVFTSQDFLTFLYTLRSYARGLRLELTLYLFQGDP 192
DB	181 TLFALADTVFTSQDFLTFLYTLRSYARGLRLELTLYLFQGDP 223
 RESULT 3 ID Q9R0C9 PRELIMINARY; PRT; 223 AA. AC Q9R0C9; DT 01-MAY-2000 (TrEMBLrel. 13, Created) DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update) DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update) DE Signal receptor (Opioid receptor, sigma 1). GN Name=Oprial; OS Rattus norvegicus (Rat). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. OX NCBI_TaxId=10116; RN [1] RP SEQUENCE FROM N.A. RC TISSUE=Brain; RA Mei J., Pasternak G.W.; RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases. [2] RN SEQUENCE FROM N.A. RP TISSUE=Prostate; RC MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899; RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., RA Brownstein M.J., Usdin T.B., Toshitoki S., Carninci P., Prange C., RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyil S.W., RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S., RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., RA Jones S.J., Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences". RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [3] RN SEQUENCE FROM N.A. RP RP RC TISSUE=Prostate; RA Strausberg R.; RL Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases. DR EMBL; AF067769; AAP08342.1; - DR EMBL; BC061978; AAA61978.1; -	

```
DR GO:0005783; C-endoplasmic reticulum; IEA.
DR GO:0000247; F-C-8 sterol isomerase activity; IEA.
DR GO:0004872; F-receptor activity; IEA.
DR GO:0006696; P-ergosterol biosynthesis; IEA.
DR InterPro: IPR006716; ERG2_Sireceptlik.
DR Pfam: PF04622; ERG2_SigmaIR; 1.
KW Receptor.
SQ
SEQUENCE 223 AA; 25270 MW; B125A0388F1FFC6E CRC64;
Query Match 92.6%; Score 940.5; DB 2; Length 223;
Best Local Similarity 80.3%; Pred. No. 2.1e-77;
Matches 179; Conservative 4; Mismatches 9; Indels 31; Gaps 1;
QY 1 MOWAVGRRWAAALLLAAVAVLTQVVMWLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
DB 1 MPWAVGRRWAWITLFTIVAVLIQAVWMLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGRHS --- 117
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGHSGRY 120
QY 118 -----GETVHGPGEATAVWGNTWMVEYGRGVIPS 149
DB 121 WAEISDTIISGTFHQWREGTTKSEVYYPGETVHGPGEATAVWGNTWMVEYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFTLTYTLRSYARGRLRLLELTLYLFGQDP 192
DB 181 TLAFALSDTIFSTQDFTLTYTLRSYARGRLRLLELTLYLFGQDP 223
RESULT 4
Q922W2 PRELIMINARY; PRT; 223 AA.
Query Match 92.6%; Score 940.5; DB 2; Length 223;
Best Local Similarity 80.3%; Pred. No. 2.1e-77;
Matches 179; Conservative 4; Mismatches 9; Indels 31; Gaps 1;
QY 1 MOWAVGRRWAAALLLAAVAVLTQVVMWLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
DB 1 MPWAVGRRWAWITLFTIVAVLIQAVWMLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGRHS --- 117
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGHSGRY 120
QY 118 -----GETVHGPGEATAVWGNTWMVEYGRGVIPS 149
DB 121 WAEISDTIISGTFHQWREGTTKSEVYYPGETVHGPGEATAVWGNTWMVEYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFTLTYTLRSYARGRLRLLELTLYLFGQDP 192
DB 181 TLAFALSDTIFSTQDFTLTYTLRSYARGRLRLLELTLYLFGQDP 223
RESULT 5
Q9JKU9 PRELIMINARY; PRT; 192 AA.
Query Match 91.0%; Score 925; DB 2; Length 192;
Best Local Similarity 89.0%; Pred. No. 4.6e-76;
Matches 170; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 1 MOWAVGRRWAAALLLAAVAVLTQVVMWLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
DB 1 MPWAVGRRWAWITLFTIVAVLIQAVWMLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGRHSGET 120
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGHSGET 120
QY 121 VVHGPGEATAVWGNTWMVEYGRGVIPSTLAFALADTVFSTQDFTLTYTLRSYARGLR 180
DB 121 VVHGPGEATALEWGPNTWMVEYGRGVIPSTLFPALADTFEFTQDYLTLTYLRAARGLR 180
QY 181 LEITTYLFGQD 191
DB 181 LEITTYLFGQD 191
RESULT 6
Q60492 PRELIMINARY; PRT; 223 AA.
Query Match 92.0%; Score 934.5; DB 2; Length 223;
Best Local Similarity 79.8%; Pred. No. 7.3e-77;
Matches 178; Conservative 4; Mismatches 10; Indels 31; Gaps 1;
QY 1 MOWAVGRRWAAALLLAAVAVLTQVVMWLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
DB 1 MPWAVGRRWAWITLFTIVAVLIQAVWMLGTSQVFQREIEAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGRHS --- 117
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALSGHSGRY 120
QY 118 -----GETVHGPGEATAVWGNTWMVEYGRGVIPS 149
```

```

RA Kempner E., Glosemann H.;
RT "Purification, molecular cloning, and expression of the mammalian
RL signal-binding site.";
RA EMBL; 266537; CAA91441.1; -.
DR GO; GO:0005783; C-endoplasmic reticulum; IEA.
DR GO; GO:0000247; F-C-8 sterol isomerase activity; IEA.
DR GO; GO:0004685; F-isomerase activity; IEA.
DR GO; GO:0004872; F-receptor activity; IEA.
DR GO; GO:0006696; F-ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_S1receptlik.
DR Pfam; PF04622; ERG2_SigmaIR; 1.
KW Isomerase; Receptor_SigmaIR; 1.
SQ SEQUENCE 223 AA; 25311 MW; 4115E51096D548F8 CRC64;

Query Match 90.4%; Score 918.5; DB 2; Length 223;
Best Local Similarity 79.4%; Pred. No. 2.1e-75;
Matches 177; Conservative 4; Mismatches 11; Indels 31; Gaps 1;

QY 1 MQWAVGRWAWAALLLAVALVTQVWLWLTQSFVFOREETIAQLAQYAGLDHLEAFSR 60
DB 1 MQWAVGRWAWAALLLAVALVTQVWLWLTQSFVFOREETIAQLAQYAGLDHLEAFSK 60
QY 61 LIVELRLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSYVLLFGTALSGRHS--- 117
DB 61 LIVELRLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSYVLLFGTALSGRHSGRY 120
QY 118 -----GETVHVGPGGATAVWGPNMTVMVEYGRGVIPS 149
DB 121 WAEISDTIISGTFHQWKEGTTKSEVFYPGETVHVGPGGATAVWGPNMTVMVEYGRGVIPS 180
QY 150 TLAFALADTVSTQDFLTFYTLRSYARGLRLLELTLYLFGQDP 192
DB 181 TLGFALADTVSTQDFLTFYTLRVYARALQLELTLYLFGQDP 223

RESULT 7
OS5242
ID O55242 PRELIMINARY; PRT; 223 AA.
AC O55242;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Type 1 sigma receptor (Oprsl protein) (Sigma 1 receptor).
GN Name=Oprsl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV, and C57BL/6; TISSUE=Kidney;
RX MEDLINE=98086386; PubMed=9425306; DOI=10.1006/bbrc.1997.7840;
RA Seth P., Leibach F.H., Ganapathy V.;
RT "Cloning and structural analysis of the cDNA and the gene encoding the
RT murine type 1 sigma receptor.";
RL Biochem. Biophys. Res. Commun. 241:535-540(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N, and FVB/N;
RC TISSUE=Liver, and Mammary tumor. WAP-TGF alpha model. 7 months old;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RC Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RC Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=97419861; PubMed=9274937; DOI=10.1016/S0014-2999(97)01064-9;
RA King M., Pan Y.X., Mei J., Chang A., Xu J., Pasternak G.W.;
RT "Enhanced kappa-opioid receptor-mediated analgesia by antisense
RT targeting the sigma receptor.";
RL Eur. J. Pharmacol. 331:R5-R6(1997).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Pan Y.-X., Xu J., Mei J.F., King M., Cang A., Wan B.-L.,
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030199; AAB97683.1; -.
DR EMBL; BC002000; AAH02000.1; -.
DR EMBL; BC019930; AAH19930.1; -.
DR EMBL; AF004927; AAC33306.1; -.
DR EMBL; AF030198; AAC39951.1; -.
DR PIR; JC5815; JCS815.
DR MGD; MGI:1195268; Oprsl.
DR GO; GO:0004872; F-receptor activity; IDA.
DR InterPro; IPR006716; ERG2_S1receptlik.
DR Pfam; PF04622; ERG2_SigmaIR; 1.
KW Receptor.
SQ SEQUENCE 223 AA; 25250 MW; 54BB2F14472B3512 CRC64;

Query Match 88.9%; Score 903.5; DB 2; Length 223;
Best Local Similarity 77.0%; Pred. No. 4.8e-74;
Matches 171; Conservative 7; Mismatches 13; Indels 31; Gaps 1;

QY 1 MQWAVGRWAWAALLLAVALVTQVWLWLTQSFVFOREETIAQLAQYAGLDHLEAFSR 60
DB 1 MPWAGRWAWITLTITIAVLQIAWLTQSFVFOREETIAQLAQYAGLDHLEAFSR 60
QY 61 LIVELRLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSYVLLFGTALSGRHS--- 117
DB 61 LIVELRLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSYVLLFGTALSGHSGRY 120
QY 118 -----GETVHVGPGGATAVWGPNMTVMVEYGRGVIPS 149
DB 121 WAEISDTIISGTFHQWKEGTTKSEVFYPGETVHVGPGGATAVWGPNMTVMVEYGRGVIPS 180
QY 150 TLAFALADTVSTQDFLTFYTLRSYARGLRLLELTLYLFGQD 191
DB 181 TLGFALADTVSTQDFLTFYTLRAYARGLRLLELTLYLFGQD 222

RESULT 8
Q8N7H3
ID Q8N7H3 PRELIMINARY; PRT; 187 AA.
AC Q8N7H3;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ25585.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Osuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK098451; BAC0307.1; -
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:000247; F:C-8 sterol isomerase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006696; P:ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_Sireceptlik.
DR Pfam; PF04622; ERG2_SigmaIR; 1.
KW Receptor.
SQ SEQUENCE 187 AA; 21013 MW; E50731B80FB98PBF CRC64;
Query Match 75.5%; Score 767.5; DB 2; Length 187;
Best Local Similarity 70.0%; Pred. No. 9.3e-62;
Matches 156; Conservative 0; Mismatches 0; Indels 67; Gaps 3;
QY 1 MQWAGRRWAWAALLLAVALTQVVMVLTQSFVFORREIAQLARQVAGLDHLEAFSR 60
DB 1 MQWAGRRWAWAALLLAVALTQVVMVLTQSFVFORREIAQLARQVAGLDHLEAFSR 38
QY 61 LIVELRLHPGHVLPDEELQWVFNAGWGMAMCLLHASLSEYVLLFGTALSGRHS--- 117
DB 39 LI-----EELQWVFNAGWGMAMCLLHASLSEYVLLFGTALSGRHSGRY 84
QY 118 -----GETVHVGGEATAVWGNTVMVVEYGRGVIPS 149
DB 85 WAEISDTISGTTFHQWREGTKSEVFPYGETVHVGGEATAVWGNTVMVVEYGRGVIPS 144
QY 150 TLAFALADTVFSTQDFLTFLYTLRSYARGLRLELTLYLFGQDP 192
DB 145 TLAFALADTVFSTQDFLTFLYTLRSYARGLRLELTLYLFGQDP 187
RESULT 9
Q6DCU6 PRELIMINARY; PRT; 221 AA.
AC Q6DCU6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Oprs1-prov protein.
GN Name=oprs1-prov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;

RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shavchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC077894; AAH7894.1; -
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:000247; F:C-8 sterol isomerase activity; IEA.
DR GO; GO:0006696; P:ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_Sireceptlik.
DR Pfam; PF04622; ERG2_SigmaIR; 1.
SQ SEQUENCE 221 AA; 24695 MW; D0D9626B03D07EA6 CRC64;
Query Match 62.6%; Score 636.5; DB 2; Length 221;
Best Local Similarity 58.0%; Pred. No. 9.1e-50;
Matches 120; Conservative 34; Mismatches 22; Indels 31; Gaps 1;
QY 12 AALLLAVALTQVVMVLTQSFVFORREIAQLARQVAGLDHLEAFSLIVELRLHPG 71
DB 9 AVLWVAGLAVLQLTRGWLSSKSYVFNREIARLAKHSGLDYEVAFSKIIVELRKHGP 68
QY 72 HVLDPDEELQWVFNAGWGMAMCLLHASLSEYVLLFGTALSGRHS----- 117
DB 69 HILQDEDLQWVFNAGWGMAMCLLHASLSEYVLLFGTAVDTGHSGRYAEISDTLSG 128
QY 118 -----GETVHVGGEATAVWGNTVMVVEYGRGVIPSTLAFALADTVF 160
DB 129 TFRQWKEGTTKSEIFYPGDTIVHVEGEATSVQSGTVMVVEYGRGFIPSTLAFALADTVF 188
QY 161 STQDFLTFLYTLRSYARGLRLELTLYL 187
DB 189 STQDFLTFLYTLRSYARGLRLELTLYL 215
RESULT 10
Q645J3 PRELIMINARY; PRT; 223 AA.
AC Q645J3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Sigma receptor.
OS Taricha granulosa (Roughskin newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae; Taricha.
OX NCBI_TaxID=8321;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Bradford C.S., Moore F.L.;
RT "Cloning, Expression, and Characterization of a Sigma Receptor from
RT the Brain of the Rough-skinned Newt, Taricha granulosa.";


```
KW Receptor.
SQ SEQUENCE 222 AA; 24940 MW; 234A53B2D4A3F3B6 CRC64;

Query Match 60.7%; Score 616.5; DB 2; Length 222;
Best Local Similarity 55.1%; Pred. No. 6e-48;
Matches 113; Conservative 32; Mismatches 29; Indels 31; Gaps 1;

QY 14 LLLAAVAALVTQVWMLWLTQSFVQREETAQARQYAGLDHSLRLVLRLLHPGHV 73
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 12 VVVGSLSTVQPIRHWMANKQVFTKEEVAKLQYAGQDHQAFKVVVLRRLPYGHI 71
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 74 LPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALGRGHS ----- 117
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 LPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALGRGHSRYWAEISDTTISGTF 131
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 118 -----GETVHGGEATAVEWGNTMMVEYGRGVIPSTLPALADTVFST 162
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 RQWKEGTTKSETYVPGDTIVHSAGEATSVQWSSGTMMVEYGRGVIPSTLGFALADTMFST 191
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 163 QDPLTLFYTLRSYARGLRLTYL 187
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 QDPLTLFYTVARVYVKGMILEASTFL 216
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
Q72653 ID Q72653 PRELIMINARY; PRT; 106 AA.
AC Q72653;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Opioid receptor, sigma 1, isoform 4.
GN Names:OPR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
EMBL; BC007839; AA07839.1; -.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0000247; F:C-8 sterol isomerase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006696; P:ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_S1receptlik.
DR Pfam; PF04622; ERG2_Sigma1r; 1.
```

```
KW Receptor.
SQ SEQUENCE 106 AA; 11992 MW; 8A4FDBFAFF6DAEEE CRC64;

Query Match 53.5%; Score 544; DB 2; Length 106;
Best Local Similarity 98.1%; Pred. No. 1.1e-41;
Matches 103; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQWAVGRWAAALLLAAVAALVTQVWMLWLTQSFVQREETAQARQYAGLDHSLR 60
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MQWAVGRWAAALLLAAVAALVTQVWMLWLTQSFVQREETAQARQYAGLDHSLR 60
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9RLJ7 ID Q9RLJ7 PRELIMINARY; PRT; 106 AA.
AC Q9RLJ7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sigma 1 receptor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RL Mei J., Pasternak G.W.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
EMBL; AF087827; AAD49439.1; -.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0000247; F:C-8 sterol isomerase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006696; P:ergosterol biosynthesis; IEA.
DR InterPro; IPR006716; ERG2_S1receptlik.
DR Pfam; PF04622; ERG2_Sigma1r; 1.
KW Receptor.
SQ SEQUENCE 106 AA; 12153 MW; 8BA4159B6A9E2A0F CRC64;

Query Match 50.0%; Score 508; DB 2; Length 106;
Best Local Similarity 89.5%; Pred. No. 2e-38;
Matches 94; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 MQWAVGRWAAALLLAAVAALVTQVWMLWLTQSFVQREETAQARQYAGLDHSLR 60
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MPWAVGRWAAITLFTIVAVLQAVWMLWLTQSFVQREETAQARQYAGLDHSLR 60
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q9C2G3 ID Q9C2G3 PRELIMINARY; PRT; 256 AA.
AC Q9C2G3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C-8 sterol isomerase erg-1.
GN Name=9G6.010;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RL Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RL Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 13:02:26 ; Search time 3089 Seconds
(without alignments)
9082.412 Million cell updates/sec

Title: US-09-823-069A-1
Perfect score: 579
Sequence: 1 atgcagtgggccgtggccg.....tctttggccaggaccttga 579

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	579	100.0	579	9 AF226604	Homo sapi
2	483	83.4	1598	9 BC007839	Homo sapi
3	483	83.4	1656	9 AX364934	Sequence
4	476	82.2	1641	9 HSU75283	Human sigma
5	476	82.2	1650	9 HSU79528	Human SR317
6	476	82.2	1655	6 CQ718427	Sequence
7	476	82.2	1679	9 BC004899	Homo sapi
8	476	82.2	1696	9 AK130502	Homo sapi
9	476	82.2	1700	6 BD076706	Human sig
10	472.8	81.7	672	9 CR457075	Homo sapi
11	438.2	75.7	579	10 AF226605	Mus muscu
12	375.2	64.8	1857	10 CPSIRECMR	C. porcellu
13	374.2	64.6	1070	10 AF087827	Rattus no
14	362.4	62.6	1117	10 AF067769	Rattus no
15	362.4	62.6	1621	10 BC061978	Rattus no
16	360.8	62.3	1590	10 AF004218	Rattus no
17	338.4	58.4	1567	10 AF030198	Mus muscu
18	338.4	58.4	1579	10 BC019930	Mus muscu
19	338.4	58.4	1595	10 BC002000	Mus muscu

20	338.4	58.4	1618	10 AF004927	Mus muscu
21	293.8	50.7	1501	9 AK098451	Homo sapi
22	235.6	40.7	34008	9 AL450283	Human DNA
23	227	39.2	1630	9 HSTIRG3	Homo sapi
24	218.2	37.7	1757	9 HSTIRG2	Homo sapi
25	203.4	35.1	1471	5 ASJ719850	Gallus ga
26	188.8	32.6	1144	5 BC077894	Xenopus la
27	185.6	32.1	531	11 BV105885	MARC 8219
28	185.8	32.1	23881	2 AC098197	Rattus no
29	185.8	32.1	242324	2 AC110351	Rattus no
30	182.6	31.5	3954	9 AY007200	Homo sapi
31	182	31.4	6973	10 AF030199	Mus muscu
32	182	31.4	168268	10 AL807796	Mouse DNA
33	171	29.5	1202	5 BC081292	Xenopus t
34	169.4	29.3	1626	5 AY726668	Taricha g
35	157.6	27.2	1175	5 BC049416	Danio rer
36	136	23.5	158408	2 BX927402	Danio rer
37	127	21.9	202	6 AR269663	Sequence
38	65.4	11.3	645	6 AB6248	Sequence
39	65.4	11.3	645	6 AR155741	Sequence
40	65.4	11.3	645	6 E66266	Genome DNA
41	65.4	11.3	110000	8 AE016819	Ermothec
42	65.4	11.3	110000	8 AE016819_01	Continuation (2 of
43	56.2	9.7	1296	6 BD179658	Highly th
44	54.2	9.4	85652	2 AC150118	Gallus ga
45	54.2	9.4	99981	2 AC150132	Gallus ga

ALIGNMENTS

RESULT 1	AF226604	Homo sapiens sigma 1 receptor beta variant mRNA, complete cds,	579 bp	mRNA	linear	PRI 18-APR-2000
LOCUS	AF226604	alternatively spliced.				
DEFINITION	AF226604					
ACCESSION	AF226604.1	GI:7582319				
VERSION						
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 579)				
AUTHORS		Wang,L.-M., Shelness,G.S., Childers,S.R., Mach,R.H. and Wheeler,K.T.				
TITLE		Sigma 1 beta receptor, an alternative splice variant of the sigma 1 receptor				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 579)				
AUTHORS		Wang,L.-M., Shelness,G.S., Childers,S.R., Mach,R.H. and Wheeler,K.T.				
TITLE		Direct Submission				
JOURNAL		Submitted (19-JAN-2000) Radiology, Wake Forest University School of Medicine, Medical Center Blvd., Winston-Salem, NC 27157, USA				
FEATURES		Location/Qualifiers				
source		1..579				
		/organism="Homo sapiens"				
		/mol_type="mRNA"				
		/db_xref="taxon:9606"				
		/cell_line="MFC-7"				
		/tissue_type="breast"				
		1..579				
		/notes="no sigma 1 binding activity, strong sigma 2 binding activity; CD region; alternatively spliced; similar to sigma 1 R1A receptor"				
		/codon_start=1				
		/product="sigma 1 receptor beta variant"				
		/protein_id="AAF6480.1"				
		/db_xref="GI:7582320"				
		/translation="MQVAGRRWAALLLAAVLTQVVMWLGTSFVFQREIIAQ LARQAGLDHDLAFSLRLVLPDEQLQVFNAGMGAMCLLASLSE YVLLFCTALGSRGSHGSETVHVGPEATVWVGNVYGRGVIPSTLAFALADTVF				
CDS						

STQDFLTLFYTLRSVARGRLRLTLYLFGQDP"									
ORIGIN									
Query Match 100.0%; Score 579; DB 9; Length 579;									
Best Local Similarity 100.0%; Pred. No. 3.2e-92;									
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGCAGTGGCGCGTGGCGCGCGGTGGCGTGGCGCGCGCTCTCTCTGGCTGTTCGACGG	60						
DB	1	ATGCAGTGGCGCGTGGCGCGCGGTGGCGTGGCGCGCGCTCTCTCTGGCTGTTCGACGG	60						
QY	61	GTGCTGACCCAGCGTCTGGCTGTGGCTGGGTAGCAGAGCTTCTCTTCCAGCGCGAA	120						
DB	61	GTGCTGACCCAGCGTCTGGCTGTGGCTGGGTAGCAGAGCTTCTCTTCCAGCGCGAA	120						
QY	121	GAGATAGCGCAGTTGGCGCGCAGTACGCTGGGCTGGACCAAGCTGGCCTTCTCTCGT	180						
DB	121	GAGATAGCGCAGTTGGCGCGCAGTACGCTGGGCTGGACCAAGCTGGCCTTCTCTCGT	180						
QY	181	CTGATCTGGAGCTGGCGCGCTGCACCCAGGCCAGTGTCTCCCGACGAGGAGCTGCAG	240						
DB	181	CTGATCTGGAGCTGGCGCGCTGCACCCAGGCCAGTGTCTCCCGACGAGGAGCTGCAG	240						
QY	241	TGSGTGTTCTGCAATGCGGSGTGGTGGATGGCGCATGTGCTTCTGCACGCTCGCTG	300						
DB	241	TGSGTGTTCTGCAATGCGGSGTGGTGGATGGCGCATGTGCTTCTGCACGCTCGCTG	300						
QY	301	TCCGAGTATGTGCTGTCTTCGCGCACCGCTTTGGGCTCCCGCGGCACCTCGGGGGAGACG	360						
DB	301	TCCGAGTATGTGCTGTCTTCGCGCACCGCTTTGGGCTCCCGCGGCACCTCGGGGGAGACG	360						
QY	361	GTAGTACACGGCGCTGTGAGCAACAGCTGTGAGTGGGGGCCCAACACATGGATGGTG	420						
DB	361	GTAGTACACGGCGCTGTGAGCAACAGCTGTGAGTGGGGGCCCAACACATGGATGGTG	420						
QY	421	GAGTACGGCGGGCGGTATCCATCCACCTTGGCTTGGCGTGGCGGACACTGTCTTTC	480						
DB	421	GAGTACGGCGGGCGGTATCCATCCACCTTGGCTTGGCGTGGCGGACACTGTCTTTC	480						
QY	481	AGCACCAGGACTTCTCACCTCTTCTATCTTCTGCTCTCTATGCTCGGGGCTCCGG	540						
DB	481	AGCACCAGGACTTCTCACCTCTTCTATCTTCTGCTCTCTATGCTCGGGGCTCCGG	540						
QY	541	CTTGAGCTCACACCTACCTCTTTGGCCAGGACCCCTTGA	579						
DB	541	CTTGAGCTCACACCTACCTCTTTGGCCAGGACCCCTTGA	579						
RESULT 2									
BC007839									
LOCUS									
DEFINITION									
Homo sapiens opioid receptor, sigma 1, transcript variant 4, mRNA									
(cDNA clone IMAGE:4124522), complete cds.									
ACCESSION									
BC007839									
VERSION									
BC007839.2									
KEYWORDS									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
AUTHORS									
1 (bases 1 to 1598)									
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,									
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,									
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,									
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,									
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,									
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,									
Schneitz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,									
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,									
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,									
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,									
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,									
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,									
Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,									
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,									
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,									
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,									
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,									
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.									
Generation and initial analysis of more than 15,000 full-length									
human and mouse cDNA sequences									
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)									
12477932									
2 (bases 1 to 1598)									
Straussberg,R.									
Direct Submission									
Submitted (11-MAY-2001) National Institutes of Health, Mammalian									
Gene Collection (MGC), Cancer Genomics Office, National Cancer									
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,									
USA									
NIH-MGC Project URL: http://mgc.nci.nih.gov									
On Dec 9, 2003 this sequence version replaced gi:14712924.									
Contact: MGC help desk									
Email: cgapbs@mail.nih.gov									
Tissue Procurement: ATCC									
cDNA Library Preparation: Rubin Laboratory									
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)									
DNA Sequencing by: National Institutes of Health Intramural									
Sequencing Center (NISC),									
Gaithersburg, Maryland;									
Web site: http://www.nisc.nih.gov/									
Contact: nisc_mgc@hgri.nih.gov									
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,									
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,									
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,									
Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,									
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,									
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,									
Young,A., Zhang,L.-H. and Green,E.D.									
Clone distribution: MGC clone distribution information can be found									
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov									
Series: IRAL Plate: 20 Row: e Column: 8									
This clone was selected for full length sequencing because it									
passed the following selection criteria: Hexamer frequency ORF									
analysis, Similarity but not identity to protein									
This clone has the following problem: The cds is short compared to									
the longest cds in the locus.									
Location/Qualifiers									
1..1598									
/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="taxon:9606"									
/clone="IMAGE:4124522"									
/tissue_type="Brain, neuroblastoma"									
/clone_lib="NIH MGC 19"									
/lab_host="DH10B-R"									
/note="Vector: pOTB7"									
1..1598									
/gene="OPRS1"									
/note="synonyms: SIGMAR1, MGC3851, SR-BP1, FLJ25585"									
/db_xref="LocusID:10280"									
/db_xref="MIM:601978"									
47..367									
/gene="OPRS1"									
/codon_start=1									
/product="opioid receptor, sigma 1, isoform 4"									
/protein_id="AAH07839.1"									

Db	113	ATGCAGTGGGGCCGTGGGGCCGGGTGGGGCCGGCTGTCTCTGGGTGTGCGACGGC	172		
Qy	61	GTGCTGACCCAGGTCGTCTGGCTCTGGCTGGGTACGACAGAGCTTCGTCTTCCAGCGCGAA	120		
Db	173	GTGCTGACCCAGGTCGTCTGGCTCTGGCTGGGTACGACAGAGCTTCGTCTTCCAGCGCGAA	232		
Qy	121	GAGATAGCGGAGTTGGGGCCGGCAGTACGCTGGGTGGACACGAGCTGGGCTTCTCTCGT	180		
Db	233	GAGATAGCGGAGTTGGGGCCGGCAGTACGCTGGGTGGACACGAGCTGGGCTTCTCTCGT	292		
Qy	181	CTGATCTGGAGCTGGGGCCGGCTGCACCCAGGCCACGCTGTGCTGCCACGACGAGGAGCTGCAG	240		
Db	293	CTGATCTGGAGCTGGGGCCGGCTGCACCCAGGCCACGCTGTGCTGCCACGAGGAGCTGCAG	352		
Qy	241	TGGGTGTTCTGGAATGCGGGTGGCTGGATGGGGCCCATGTGCCCTTCTGACGCTCGCTG	300		
Db	353	TGGGTGTTCTGGAATGCGGGTGGCTGGATGGGGCCCATGTGCCCTTCTGACGCTCGCTG	412		
Qy	301	TCCGAGTATGTGCTCTTTCGGGCACCGGCTTGGGGCTCCCGCGGCCACTCG	351		
Db	413	TCCGAGTATGTGCTCTTTCGGGCACCGGCTTGGGGCTCCCGCGGCCACTCGGGCGCTAC	472		
Qy	352	-----	351		
Db	473	TGGGCTGAGATCTCCGATACCATCATCTTGGGCACCTTCCACAGTGGAGAGGGGCACC	532		
Qy	352	-----GGGGAGACGGTAGTACACGGGGCTGGTAGGGCAACA	387		
Db	533	ACCAAAAGTGGGTCTTCTACCCAGGGGAGACGGTAGTACACGGGGCTGGTAGGGCAACA	592		
Qy	388	GCTGTGAGTGGGGCCCAACACATGATGTTGGAGTACGGCCGGGGCGTCAATCCATCC	447		
Db	593	GCTGTGAGTGGGGCCCAACACATGATGTTGGAGTACGGCCGGGGCGTCAATCCATCC	652		
Qy	448	ACCTTGGCTTCGGGCTGGCCGACACATGCTCTTACGACACCCAGGACTTCTCACCCCTCTTC	507		
Db	653	ACCTTGGCTTCGGGCTGGCCGACACATGCTCTTACGACACCCAGGACTTCTCACCCCTCTTC	712		
Qy	508	TATACTTTCGCTCTCTATGCTCGGGGCTTCGGGCTTGAAGTTCACCACTACCTCTTTGGC	567		
Db	713	TATACTTTCGCTCTCTATGCTCGGGGCTTCGGGCTTGAAGTTCACCACTACCTCTTTGGC	772		
Qy	568	CAGGACCCCTTGA 579			
Db	773	CAGGACCCCTTGA 784			
RESULT 9	BD076706	1700 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	Human sigma receptor.				
DEFINITION	Human sigma receptor.				
ACCESSION	BD076706				
VERSION	BD076706.1	GI:22622309			
KEYWORDS	JP 2001515714-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Hillman, J.L., Corley, N.C. and Shah, P.				
JOURNAL	Human sigma receptor				
COMMENT	Patent: JP 2001515714-A 1 25-SEP-2001;				
	INCYTE PHARMACEUTICALS INC				
	OS Homo sapiens (human)				
	PN JP 2001515714-A/1				
	PD 25-SEP-2001				
	PR 11-SEP-1998	JP 2000510844			
	PR 12-SEP-1997	US 08/928612			
	PI JENNIFER L HILLMAN, NEIL C CORLEY, PURVI SHAH				
	PC CL2N15/09, A61K38/00, A61K39/395, A61K45/00, A61P1/04, A61P3/00, PC				
	A61P3/00,				
	PC A61P10, A61P5/14, A61P9/02, A61P11/06, A61P13/02, A61P13/12, PC				
	A61P17/02.				

Medicine, Medical Center Blvd., Winston-Salem, NC 27157, USA

1. .579
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Hej"
/db_xref="taxon:10090"
/tissue_type="breast"

1. .579
/notes="no sigma 1 binding activity, strong sigma 2 binding activity; CD region; alternatively spliced"
/codon_start=1
/product="sigma 1 receptor beta variant"
/protein_id="AAF64281.1"
/db_xref="GI:7582322"
/translations="MPWAGRRWAWITLITIAVLIOAWLWLTQNFVFEETIAQ
LARQAGLDHLEAFSLIVELRLHFGVLPDEELQWVFNAGWGMAGMCLHSLASE
YVLLFGTALGSHGSHGETVHGFGEATALEWGPNTWVYGRGVIPTSLFFALADTFE
GTQDYLTFLPYTLRAYARGLRLLELTLYLFGQDS"

ORIGIN

Query Match 75.7%; Score 438.2; DB 10; Length 579;
Best Local Similarity 84.8%; Pred. No. 1.9e-67;
Matches 491; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 ATGCAGTGGGCGCTGGGCGGCGTGGGCGTGGGCGGCTGCTCTGGCTGTCGCGAG 60
DB 1 ATCCGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60

QY 61 GTGCTGACCCAGGCTGCTGCTGGCTGCTGGCTGGGCTACGACAGAGTTCGCTTCCACGCGAA 120
DB 61 GTGCTGATCCAGCGCGCTGCTGCTGGCTGGGCGCTCAAACTTCGCTTCTCTAGAGAA 120

QY 121 GAGATACGCGAGTTGGGCGGCGAGTACGCTGGGCTGGGCTGACACAGAGCTGCTCTCGT 180
DB 121 GAAATACGCGAGTTGCTCGACAGTATGCGGGGCTGGACCATGAGCTTGCCTTCTCGG 180

QY 181 CTGATCGTGAGCTGGGCGGCTGCACCCAGGCGCAGCTGCTGCCGACGAGGAGCTGCAG 240
DB 181 CTGATCGTGAGCTGGGCGGCTGCACCCAGGCGCAGCTGCTGCCGATGAGGAGCTGCAG 240

QY 241 TGGGTGTTCTGTAATGGGCGTGGCTGATGGGCGGCGCATGTGCTTCTGACGCGCTGCTG 300
DB 241 TGGGTATTTGTGAACGGGCGGCTGATGGGCGGCGCATGTGATTTCTGACGCGCTGCTG 300

QY 301 TCCGAGTATGTGCTGCTTTCGCGACCCGCTTGGGCTCCCGGCGCCTATCGGGGAGAGC 360
DB 301 TCTGAGTACGTGCTCTTTCGCGACCCGCTTGGGCTCCCGGCGCCTATCGGGGAGACA 360

QY 361 GTAGTACACGGGCTGTGAGGACACAGCTGTGAGTGGGCGGCGCAACACATGATGATG 420
DB 361 GTTGTACACGGGCTTGAGAGAACAGGCTGTGAGTGGGCGGCGCAACACATGATGATG 420

QY 421 GAGTACGGGCGGCGCTCATCCCATCCAGCTGCGCTTCCGCTTGGCGGCGACACTGTCTTC 480
DB 421 GAGTACGGGCGGCGGTATTTCCGCTTACCTGCTTCTTTCGATAGCGGACACTTCTTC 480

QY 481 AGCACCCAGACTTCTCACCTTCTTATACCTTCTGCTCTTATGCTCGGGGCTCCGG 540
DB 481 GGCACCCAGACTTACCTTACACTTCTTATACCTTCTGCGGCTATGCCGGGCGCTCCGG 540

QY 541 CTTGAGCTTACCACCTTCTTGGCGAGGACCTTGA 579
DB 541 CTTGAGCTTACCACCTTCTTGGCGAGACTTCTGA 579

RESULT 12
CPS1RECMR 1857 bp mRNA linear ROD 07-OCT-1996
LOCUS C.porcillus mRNA for signal-receptor.
DEFINITION Z66537
ACCESSION Z66537.1 GI:1403299
VERSION
KEYWORDS signal-receptor; sterol-isomerase.

SOURCE Cavia porcellus (domestic guinea pig)
ORGANISM Cavia porcellus
REFERENCE 1 (bases 1 to 1857)
AUTHORS Hanner,M., Moebius,F., Flandorfer,A., Knaus,H.G., Striessnig,J., Kempner,E. and Glossmann,H.
TITLE Purification, molecular cloning, and expression of the mammalian signal-binding site
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (15), 8072-8077 (1996)
MEDLINE 96333947
PUBMED 8755605
REFERENCE 2 (bases 1 to 1857)
AUTHORS Hanner,M.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-1995) Hanner M., Universitaet Innsbruck, Biochemische Pharmakologie, Peter-Mayr-Str. 1, Innsbruck, Austria, A-6020

Location/Qualifiers
1. .1857
/organism="Cavia porcellus"
/mol_type="mRNA"
/strain="Pirbright white"
/db_xref="taxon:10141"
/clone="lambda-Gp8"
/sex="female"
/tissue_type="liver"
/clone_lib="lambda gt10"
378..1049

CDS
1. .1857
/product="signal-receptor (sterol-isomerase like protein)"
/protein_id="CAA91441.1"
/db_xref="GI:1403300"
/db_xref="GOA:Q60492"
/translation="MQWAGRRWLWVLFPAVAVLTQIWLWLTQNFVFEETIAQ
LARQAGLDHLEAFSLIVELRLHFPVLPDEELQWVFNAGWGMAGMCLHSLASE
YVLLFGTALGSPRHSRYWAEISDTTISGTFHQWREGTTKSEYFPGETVHGFGEAT
AWEWGPNTWVYGRGVIPTSLGFLADTVFSTQDPLTLFYTLRVARALQLELTLYL
FQQDP"
1834..1839

polyA_site
ORIGIN

Query Match 64.8%; Score 375.2; DB 10; Length 1857;
Best Local Similarity 76.8%; Pred. No. 1.8e-56;
Matches 516; Conservative 0; Mismatches 63; Indels 93; Gaps 1;

QY 1 ATGCAGTGGGCGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 60
DB 378 ATGCAGTGGGCGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 437

QY 61 GTGCTGACCCAGTTCGCTGCTGGCTGCTGGCTGGGCTACGACAGAGCTTCGCTTCCAGCGGAA 120
DB 438 GTTCTAACCAATTCGTGTGGCTCTGGCTGGGCGACGAGAACTTCGCTTCCAGCGGAA 497

QY 121 GAGATACGCGAGTTGGCGGCGAGTACGCTGGGCTGGACCAACGAGCTGGGCTTCTCTCGT 180
DB 498 GAAATCGGCGAGCTGGGCGGCGAGTACGCTGGCTGGACCAACGAGCTTCTCTAAG 557

QY 181 CTGATCGTGAGAGTGGGCGGCTGCACCCAGGCGACGCTGCTGCCGACGAGAGCTGCAG 240
DB 558 CTGATCGTGAGAGTGGGCGGCTGCACCCGCGTACAGTACTGCCCGATGAGGAGCTGCAG 617

QY 241 TGGGTGTTCTGTAATGGGCGGCTGATGGGCGGCGCATGTGCTTCTGACGCTCGCTG 300
DB 618 TGGGTGTTCTGTAACGGGCGGCTGATGGGCGGCGCATGTGCTTCTGACATGCCCTG 677

QY 301 TCCGAGTATGCTGCTTCTTCGCGACCGCTTGGGCTCCCGCGGCGCACTCG----- 351
DB 678 TCCGAGTATGCTGCTTCTTCGCGACCGCTTGGGCTCACCGGCGCACTCGGGCGGCTAT 737
QY 352 ----- 351

```
Db      738 TGGGCTGAGATCTCCGACACCATATATCTCTGGCACTTTCCACGAGTGAGAGAGGGCACC 797
Qy      352 -----GGGAGACGGTAGTACACGGGCCCTGGTGAGGCAACA 387
Db      798 ACCAAAAGTAGGCTCTTATCTACAGGGGAGACAGTGGTGCACGGGCCCTGGTGAAGCAACG 857
Qy      388 GCTGTGGAGTGGGGCCAAACACATGGATGGTGGAGTACGGCCGGGGGCTCATCCCATCC 447
Db      858 GCTGTGGAAATGGGGACCAAAACACATGGATGGTGGAGTACGGCCGGGGGCTCATCCCGTCT 917
Qy      448 ACCCTGGGCTTCGGCCTGGCCGACACTCTCTTCAGCACCCAGGACTTCTCTCACCCCTCTTC 507
Db      918 ACCCTGGGATTTGCACTGGCTGACACTGTCTTCAGCACCCAGGACTTCTCTCACCCCTCTTC 977
Qy      508 TATACCTTCGCTCCTATGCTCGGGGCTTCGGGCTTGAGCTCACCACTACCTCTTTGGC 567
Db      978 TATACCTTCGAGTCTATGCGCGGGCCCTCCAGCTTGAACCTCACCACTACCTCTTCGGC 1037
Qy      568 CAGGACCCCTGA 579
Db      1038 CAAGACCCCTGA 1049

RESULT 13
AF087827 AF087827 1070 bp mRNA linear ROD 16-AUG-1999
DEFINITION Rattus norvegicus sigma 1 receptor mRNA, alternatively spliced,
complete cds.
ACCESSION AF087827
VERSION AF087827.1 GI:57333129
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 1070)
AUTHORS Mei, J. and Pasternak, G.W.
TITLE Molecular Cloning and Characterization of Rat Sigma 1 Receptor and
Its Splice Isoform
JOURNAL Unpublished
AUTHORS Mei, J. and Pasternak, G.W.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1998) Neurology, Memorial Sloan-Kettering Cancer
Center, 1275 York Ave., New York, NY 10021, USA
FEATURES
source
1..1070
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
91..411
/note="alternatively spliced; deletion results in a
truncated protein when compared with the product encoded
in GenBank Accession Number AF067769"
CDS
/product="sigma 1 receptor"
/protein_id="AAD49439.1"
/db_xref="GI:57333130"
/translation="MPWAVGRWAWITPLTIVAVLIQAVMLGTSFVFORREIAQ
LARQYAGLDHLEAFSLRVELRHPGHVLPDEELQWVFNAGWGMGACLLHLSLE
TILG"

ORIGIN
Query Match 64.6%; Score 374.2; DB 10; Length 1070;
Best Local Similarity 77.8%; Pred. No. 3 1e-56;
Matches 486; Conservative 0; Mismatches 93; Indels 46; Gaps 1;

Qy      1 ATGCAGTGGCGCGTGGCGCGGCTGGCGCGGCTCTCTCTGCTCTCGAGCG 60
Db      91 ATGCCGTGGCTGTGGCGCGGCTGGCGCGGCTGGCGCGGCTCTCTCTGCTCTGAGCG 150
Qy      61 GTGCTGACCCAGGTGCTGTGGCTCTGGCTGTGGGTACGACGAGCTTCGTCTTCCAGCGCGNA 120
```

```
Db      151 GTGCTGATCCAGGCGGCTCTGGCTGGCTGGTACTCAGAGCTTCGTCTTCCAGAGAGAA 210
Qy      121 GAGTAGGCGAGTTGGGCGCGCAGTACGCTGGGCTGGACCAAGAGAGCTGGCTTCTCTGTT 180
Db      211 GAGTAGCTCAGCTTGTCTCGACAGTACGCGAGGGTGGACCAATGATGCTTCTTCTCGG 270
Qy      181 CTGATCGTGGAGCTGCGGCGGCTGCACCCAGGCCACAGTGTGCTGCCGACGAGAGAGCTGCAG 240
Db      271 CTGATCGTGGAGCTGCGGAGGCTGCACCCAGGCCACAGTGTGCTGCCGATGAGGAGCTGCAG 330
Qy      241 TGGGTGTTTCGTGAATGGCGGCTGGATGGAGGCGCCATGTGCTTCTGCACGCTTCGCTG 300
Db      331 TGGGTGTTTGTGAACGGGCGGCTGGATGGAGGCGCCATGTGCTTCTTTCACGCTTCGCTG 390
Qy      301 TCCGAGTATGTGCTGCTCT----- 319
Db      391 TCTGAGACGATACCTGGGCTGAGATTTCCAGATACCATCATCTCTGGCACTTTTCCACAGGTG 450
Qy      320 -----TCGGCACCGCTTGGGCTCCCGCGGCACCTCGGGGGAGAGCGGTAGTACAGGGCC 374
Db      451 GAGAGAGGCGCACCAAAAAGTGAGGCTTATTAACCAGAGAGAGCGTTGTCCATGGACC 510
Qy      375 TGGTACAGCAACAGCTGTGGAGTGGGGGCCAAAACACATGGATGGTGGAGTACGGCCGGG 434
Db      511 TGGAGAGCAACAGCTGTGGAGTGGGACCAAAACACAGTGGATGGTGGAGTATGGCCGGG 570
Qy      435 CGTCATCCATCCACCTGGGCTTCGCGCTGGCGGACACTGTCTTTCAGCAGCCAGACTT 494
Db      571 TGTATTTCGCTACACCTGGCATTTGCACCTTAGTGACACTATTTTTCAGACCCAGGACTT 630
Qy      495 CCTCACCTCTTCTATCTCTTCGCTCTGCTCGGGGCTTCGGGCTTGAGCTCACCAC 554
Db      631 CCTCACCTCTTCTATACCTTCGCGCTATGCCGGGCTATGCCGGGCTTCGGCTTGAGCTCACCAC 690
Qy      555 CTACTCTTTTGGCCAGGACCCCTTGA 579
Db      691 CTACTCTTTTGGCCAGGACCCCTGA 715

RESULT 14
AF067769 AF067769 1117 bp mRNA linear ROD 03-JUN-2004
DEFINITION Rattus norvegicus sigma 1 receptor mRNA, complete cds.
ACCESSION AF067769
VERSION AF067769.1 GI:6434854
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 1117)
AUTHORS Mei, J. and Pasternak, G.W.
TITLE Molecular cloning and pharmacological characterization of the rat
sigma 1 receptor
JOURNAL Biochem. Pharmacol. 62 (3), 349-355 (2001)
MEDLINE 21328659
PUBMED 11434908
REFERENCE 2 (bases 1 to 1117)
AUTHORS Mei, J. and Pasternak, G.W.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Neurology, Memorial Sloan-Kettering Cancer
Center, 1275 York Avenue, New York, NY 10021, USA
FEATURES
source
1..1117
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain"
/note="Cesarian derived strain"
91..762
/codon_start=1
CDS
```

		/product="sigma receptor" /protein_id="AAF08342.1" /db_xref="GI:6434855" /translation="MPWAVGRWAMITPLTIVAVLIQAVWLGTQSFVFOREBIAQ LARQVAGLDHELAFLRLIVELRLHPGHVLPDEELQWVFNAGGWMGACMLLHSLSE VYLLFTGALSHGHSGRYWAETSDTIIISGTFHQWREGTTKSEVYVPGETVHVGPEAT AVEWGNTWMVEYGRGVIPTSLAFALSDTIFSTQDFLTFLYTLRAYARGRLRLATLYL FQDQp"	
ORIGIN			
Query Match	62.6%; Score 362.4; DB 10; Length 1117;		
Best Local Similarity	75.6%; Pred. No. 3.6e-54;		
Matches 508; Conservative	0; Mismatches 71; Indels 93; Gaps 1;		
QY	1	ATGCAGTGGCGCGTGGCGCGCGGTGGCGGTGGCGCGCGCTCTCTGCTGTGCGAGCG 60	
Db	91	ATGCCGTGGGCTGTGGCGCGCGGTGGCGATGAGTACACCTGTCTTGACTATTGGCGG 150	
QY	61	GTGCTGACCCAGGTCTGTGGCTCTGGCTGGGTGACGAGCTTTCCTTCCAGCGCGAA 120	
Db	151	GTGCTGATCAGCGCGTGTGGCTGTGGCTGGGTGACTCAGAGCTTCTCTCCAGAGAA 210	
QY	121	GAGATAGCGAGTTGGCGCGGACGTACGCTGGGCTGGACCAAGAGCTGCGCTTCTCTGT 180	
Db	211	GAGATAGCTCAGCTTCTCGACAGTACGCGAGGCTGGACCATGAGCTGGGCTTCTCTCG 270	
QY	181	CTGATCGTGAGTGGCGCGCTGACCCAGGCCACGTGTGCTGCCCGACGAGAGCTGCAG 240	
Db	271	CTGATCGTGAGTGGCGAGGCTGCACCCAGGCCACGTGTGCTGCCGATGAGGAGCTGCAG 330	
QY	241	TGGGTCTGCTGAATCGGCGTGGCTGATGGCGGCGCATGTGCTTCTGCAAGCTCGCTG 300	
Db	331	TGGGTGTTGTGAACGGCGCGGCTGGATGGCGGCGCATGTGTCTTTTGACGCGCTCGTG 390	
QY	301	TCCGAGTATGTGCTCTTTCGCGACCGCTTTCGGGCTCCGCGGCCACTCG----- 351	
Db	391	TCTGAGTACGTGTGCTCTTTCGCGACCGCTTCGCGCTCCCATGSCCATTCGGGACGATC 450	
QY	352	----- 351	
Db	451	TGGGCTGAGATTTCAGATACCATCATCTCTGGCACATTTTCCACCAGTGGAGAGGCGACC 510	
QY	352	-----GGGGACCGTAGTACACGGCGCTGTGAGGCAACA 387	
Db	511	ACAAAAAGTGAGTCTATTACCCAGGAGAGACGGTGTCTCTGACCTGAGAGGCAACA 570	
QY	388	GCTGTGAGTGGGGGCAACACATGATGGTGGAGTACGGCGGGCGGTCTATCCCATCC 447	
Db	571	GCTGTGAGTGGGGACCAACACGTGGATGTGGAGTATGCGCGGGGTGTATTCCGTCT 630	
QY	448	ACCTTGCGCTTGGCTGGCGGACACATGTCTTTCAGACCCAGAGCTTCTCACCCCTTTC 507	
Db	631	ACCTTGCGATTTGCACCTTAGTGACATATTTTCAGCACCCAGGACTTCTCTCACCCCTTTC 690	
QY	508	TATACCTTTCGCTTCTATGCTCGGGCGCTCCGCTTGAGCTCACCACTACCTCTTGGC 567	
Db	691	TATACCTTTCGCGCTTATGCTCGGGCGCTCCGCTTGAGCTCACCACTACCTCTTGGC 750	
QY	568	CAGGACCCCTTGA 579	
Db	751	CAAGACCCCTTGA 762	
RESULT 15			
BC061978			
LOCUS			
DEFINITION	BC061978	1621 bp mRNA linear ROD 30-JUN-2004	
		Rattus norvegicus opioid receptor, sigma 1, mRNA (cDNA clone	
		MGC:72418 IMAGE:5623693), complete cds.	
ACCESSION	BC061978		
VERSION	BC061978.1	GI:38541100	
KEYWORDS			
SOURCE			
ORGANISM			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 1621)

Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Workley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1621)

Strausberg,R.

Direct Submission

Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue procurement: Dr. Maarten Bosland, NYU

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 137 Row: O Column: 24

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13591887.

Location/Qualifiers

1. .1621

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:72418 IMAGE:5623693"

/tissue_type="Prostate, normal lateral, 14 mo. old male rat"

/clone_lib="NCI CGAP_Pr33"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

1. .1621

/gene="Opr1"

/db_xref="LocusID:29336"

/db_xref="RGD:68364"

FEATURES	
source	
gene	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 10:54:55 ; Search time 479 Seconds
(without alignments)

7155.595 Million cell updates/sec

Title: US-09-823-069A-1

Perfect score: 579

Sequence: 1 atgcagtgccgcgtggcg.....tcttggccagacccttga 579

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	579	100.0	579	4	AAS15569 DNA encod
2	483	83.4	1656	6	ABL39775 Human NS
3	476	82.2	1650	13	ACN38862
4	476	82.2	1655	12	ADQ87384
5	476	82.2	1700	2	AAK39478
6	473	81.7	696	6	AAH77235
7	438.2	75.7	579	4	AAS15570
8	387	66.8	1535	3	AAF21650
9	360.8	62.3	1590	10	ADB58553
10	158.4	27.4	442	9	ACH50720
11	127	21.9	202	10	ACA55628
12	127	21.9	202	12	AD155424
13	93	16.1	498	9	ACH45312
14	65.4	11.3	645	2	ADR02215
15	51.4	8.9	669	10	ADK65614
16	51.4	8.9	669	13	ADSA7159
17	51.4	8.9	1508	2	AAT11777
18	49.6	8.5	1052	13	ADT44395
19	49	8.5	633	6	ABA94360
20	49	8.5	636	6	ABV78180

21	49	8.5	636	6	AB235756	Abz35756 Human FGF
22	49	8.5	636	6	ABK86901	Abk86901 Human fib
23	49	8.5	636	6	ABX09999	Abx09999 Human FGF
24	49	8.5	636	6	ABL91721	AbL91721 Human pol
25	49	8.5	701	6	ABA94359	AbA94359 Human FGF
26	49	8.5	1330	4	AA94359	AA94359 Human fib
27	48.6	8.4	3018	8	AAL61203	AaL61203 Actinosyn
28	48.6	8.4	82746	8	AAL61224	AaL61224 Actinosyn
29	47.4	8.2	633	5	AAF62049	AaF62049 Human fib
30	47.4	8.2	633	6	ABN84523	Abn84523 Human fib
31	47.4	8.2	633	6	ABK13368	Abk13368 Human CDN
32	47.4	8.2	633	8	ACF57059	AcF57059 Human fib
33	47.4	8.2	633	10	ABQ83437	AbQ83437 Human fib
34	47.4	8.2	633	12	ADE39174	AdE39174 Human FGF
35	47.4	8.2	633	12	ADE39206	AdE39206 Human FGF
36	47.4	8.2	633	12	ADM57640	AdM57640 Human fib
37	47.4	8.2	636	3	AAA75630	Aaa75630 Nucleotid
38	47.4	8.2	636	5	AAO3277	AaO3277 Human CDN
39	47.4	8.2	636	6	AAD39242	AaD39242 Human FGF
40	47.4	8.2	636	8	ACF57077	AcF57077 Human FGF
41	47.4	8.2	636	12	ABE39192	AbE39192 Human FGF
42	47.4	8.2	814	6	ABK13369	Abk13369 Human DNA
43	47.4	8.2	814	8	ACF57060	AcF57060 Human fib
44	47.4	8.2	814	10	ABQ83442	AbQ83442 Human fib
45	46.8	8.1	855	13	ADS49360	AdS49360 Bacterial

ALIGNMENTS

RESULT 1

AAS15569

ID AAS15569 standard; cDNA; 579 BP.

XX AAS15569;

XX 29-JAN-2002 (first entry)

XX DNA encoding human sigma 1 receptor splice variant, sigma1beta.

XX Human; sigma 1 receptor; sigma1beta; tumour imaging; cancer.

KW cell proliferation disorder; tumour; diagnostic; cytostatic; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..579

FT /*tag= a

FT /product= "Sigma 1 receptor splice variant sigma1beta"

XX WO200174297-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010650.

XX 31-MAR-2000; 2000US-0193694P.

XX (UYWA-) UNIV WAKE FOREST.

XX Wheeler KT, Mach RH, Childers S, Shelness G, Wang L;

XX WPI; 2001-662943/76.

XX P-PSDB; AAU09154.

XX Novel isolated polynucleotide encoding sigma1beta receptor useful in

XX screening assay to identify ligands specific for the sigma1beta receptor

XX for tumor imaging, diagnostic and treatment methods.

XX Claim 1; Fig 2; 56pp; English.

XX The invention relates to an isolated polynucleotide (I) encoding sigma 1

XX receptor splice variant isolated from mouse adenocarcinoma cells and

XX		10-JUN-1999	(first entry)	
XX		Human sigma receptor (SIGR) polypeptide encoding DNA.		
XX		Sigma receptor; SIGR; vesicle-trafficking disorder; gastrointestinal;		
XX	KW	inflammatory; immunological disorder; AIDS; arthritis; infection; cancer;		
KW	KW	neoplastic disorder; ss.		
XX		Homo sapiens.		
OS		US5863766-A.		
PN	PN	26-JAN-1999.		
XX	PD	12-SEP-1997; 97US-00928612.		
XX	PP	12-SEP-1997; 97US-00928612.		
PR	PR	(INCY-) INCYTE PHARM INC.		
PA	PA	Corley NC, Hillman JL, Shah P;		
PI	PI	WPI: 1999-253089/21.		
DR	DR	P-PSDB; AAY03755.		
XX		New DNA encoding human sigma receptor polypeptide - useful for treating		
PT	PT	or preventing vesicle-trafficking, immunological and neoplastic		
PT	PT	disorders.		
XX		Claim 4; Fig 1A-D; 28pp; English.		
XX		This DNA encodes a human sigma receptor (SIGR) polypeptide. Host cells		
CC	CC	containing a vector comprising the SIGR nucleic acid are used for the		
CC	CC	recombinant expression of the protein. Recombinant SIGR can be used to		
CC	CC	treat or prevent vesicle-trafficking disorders, e.g cystic fibrosis.		
CC	CC	glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes		
CC	CC	insipidus, hyper-and hypoglycemia, Grave's disease, goiter, Cushing's		
CC	CC	disease, and Addison's disease; gastrointestinal disorders (ulcerative		
CC	CC	colitis, gastric and duodenal ulcers); and other conditions associated		
CC	CC	with abnormal vesicle trafficking (e.g. hay fever, urticaria (hives),		
CC	CC	inflammatory bowel disease, myasthenia gravis, Chediak-Higashi syndrome,		
CC	CC	systemic lupus erythematosus, toxic shock syndrome, and traumatic tissue		
CC	CC	damage). SIGR can be used to treat or prevent immunological disorders		
CC	CC	such as AIDS, adult respiratory distress syndrome, allergies, anemia,		
CC	CC	asthma, atherosclerosis, bronchitis, cholecystitis, atopic dermatitis,		
CC	CC	Crohn's disease, dermatomyositis, diabetes mellitus, glomerulonephritis,		
CC	CC	gout, hypereosinophilia, lupus erythematosus, multiple sclerosis,		
CC	CC	myasthenia gravis, myocardial or pericardial inflammation, rheumatoid		
CC	CC	arthritis, osteoarthritis, osteoporosis, pancreatitis, polymyositis,		
CC	CC	scleroderma, Sjogren's syndrome, Werner syndrome, autoimmune thyroiditis;		
CC	CC	complications of cancer hemodialysis, and extracorporeal circulation; and		
CC	CC	parasitic, viral, bacterial, fungal, helminth, and protozoal infections.		
CC	CC	SIGR may also be used to treat or prevent neoplastic disorders e.g.		
CC	CC	adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma,		
CC	CC	teratocarcinoma, and cancers of the adrenal gland, bladder, ganglia,		
CC	CC	gastrointestinal tract, heart, breast, cervix, liver, lung, muscle,		
CC	CC	ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin,		
CC	CC	spleen, testis, thymus, thyroid, and uterus		
XX		Sequence 1700 BP; 342 A; 525 C; 479 G; 354 T; 0 U; 0 Other;		
SQ		Query Match 82.2%; Score 476; DB 2; Length 1700;		
		Best Local Similarity 86.2%; Pred. No. 6.6e-96;		
		Matches 579; Conservative 0; Mismatches 0; Indels 93; Gaps 1		
QY	1	ATGCAGTGGGCGGTGGGCGGGTGGGCGTGGGCGGCTGCTCCTGGCTTCGACGCG 60		
Dd	116	ATGCAGTGGGCGGTGGGCGGGTGGGCGGTGGGCGGCTGCTCCTGGCTTCGACGCG 175		
QY	61	GTGCTACCACAGTGTGTGGCTGTGGCTGGGTACGACAGCTTCGTTCCACGCCGAA 120		
Dd	176	GTGCTACCACAGTGTGTGGCTGTGGCTGGGTACGACAGCTTCGTTCCACGCCGAA 235		

Matches 491; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 ATGCACTGGCGCGTGGCGCGGCGTGGCGCGCGCTGCTGCTGCTGCTGCGAGCG 60
DB 1 ATGCGGTGGCGCGCGGAGCGCGGTGGCGATGATCACCCTGATCTGACTATTATTCGCA 60

QY 61 GTGCTGACCCAGGCTGCTGCGCTCTGCGCTGGGTACGACAGAGCTTCGTTCCAGCGCGAA 120
DB 61 GTGCTGATCCAGCGCGCTGGTTGTGCTGGGCACTCAAACCTTCGTTCTCTAGAGAA 120

QY 121 GAGATAGCGAGTTGGCGGCGAGTAGCGCTGGCTGGACACAGAGCTGGCTTCCTCGT 180
DB 121 GAAATAGCGAGCTTGTCTGACAGTATGCGGGCTGGACCATGAGCTTGCCTTCTCTCG 180

QY 181 CTGATCGTGGAGCTGGCGGCTGACCCAGGCGCAGTGTGCTGCCACGAGGAGCTGCAG 240
DB 181 CTGATCGTGGAGCTGGCGGCTGACCCAGGCGCAGTGTGCTGCCGATGAGGAGCTGCAG 240

QY 241 TGGGTGTTGTAATGCGGTGGCTGATGGCGCCCATGTGCTTCTGACGCGCTTCGCTG 300
DB 241 TGGGTATTGTGAACGCGGCGGCTGGATGGCGCCCATGTGTTCTGACGCGCTTCGCTG 300

QY 301 TCGAGTATGTGCTCTTCGCGACCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCG 360
DB 301 TCTGAGTACGTGCTCTTCGCGACCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCG 360

QY 361 GTAGTACACGGGCTGTGAGGCAACAGCTGTGAGTGGGGGCGCAACACATGATGCTG 420
DB 361 GTTGTACACGGGCTGTGAGGCAACAGCTGTGAGTGGGGGCGCAACACATGATGCTG 420

QY 421 GAGTACGGCGGGCGCTCATCCCATCCCGCTTCGCGCTTCGCGCTTCGCGCTTCGCTTC 480
DB 421 GAGTACGGCGGGGTGTTATTCGCTTACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 480

QY 481 AGCACCCAGGACTTCCTCACCTCTTCTATPACTTTCGCTTCCTATGCTCGGGGCTCCGG 540
DB 481 GGCACCCAGGACTTACCTCACCTCTTCTATPACTTTCGCGCTTCGCGCTTCGCGCTTCG 540

QY 541 CTTGAGCTACCACTTACCTCTTTGGCGGAGCCCTTGA 579
DB 541 CTTGAGCTTACCCTTACCTCTTTGGCGGAGACTCTCTGA 579

RESULT 8
AAF21650
ID AAF21650 standard; DNA; 1535 BP.

XX AAF21650;

XX 27-MAR-2001 (first entry)

DE Human breast and ovarian cancer associated antigen gene SEQ ID 37.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW notropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnarary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.

XX Homo sapiens.

OS WO200055173-A1.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US005881.

PF 12-MAR-1999; 99US-0124270P.

PR (HUMA-) HUMAN GENOME SCI INC.

PA

XX PI Rosen CA, Ruben SM;
XX WPI: 2000-611515/58.
DR P-PSDB; AAB58747.
XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention, treatment
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
PT neurological diseases.
XX Claim 1; Page 508; 1299pp; English.
PS Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive; notropic;
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC antiinflammatory; antitumor; vulnarary; anticonvulsant; antibacterial;
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC protein sequences are used in the diagnosis of cancer, particularly
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC and agonists may also be used in the diagnosis, prevention and treatment
CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemias; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
CC infectious diseases
XX Sequence 1535 BP; 349 A; 457 C; 402 G; 325 T; 0 U; 2 Other;
SQ

Query Match 66.8%; Score 387; DB 3; Length 1535;
Best Local Similarity 84.0%; Pred. No. 3.4e-76;
Matches 490; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 90 GGGTAGCGAGCTTCGCTTCAGCGCGAAGAGATAGCGAGCTTCGGCGGCGAGTACGC 149
DB 8 GGGTAGCGAGCTTCGCTTCAGCGCGAAGAGATAGCGAGCTTCGGCGGCGAGTACGC 67

QY 150 TGGGCTGGACCGAGCTTCGCTTCGCTTCGATCGTGGAGCTGCGGCGGCTGCACCC 209
DB 68 TGGGCTGGACCGAGCTTCGCTTCGCTTCGATCGTGGAGCTGCGGCGGCTGCACCC 127

QY 210 AGGCCACGTGCTGCCGACGAGAGCTGCAGTGGGTGTTGCTGAATGCGGCTGGAT 269
DB 128 AGGCCACGTGCTGCCGACGAGAGCTGCAGTGGGTGTTGCTGAATGCGGCTGGAT 187

QY 270 GGGCGCATGTCCTTCGACGCGCTGCTGCGAGTATGCTGCTTCGCGCACCCG 329
DB 188 GGGCGCATGTCCTTCGACGCGCTGCTGCTGCGAGTATGCTGCTTCGCGCACCCG 247

QY 330 CTTGGGCTCCCGCGGCACTCG----- 351
DB 248 CTTGGGCTCCCGCGGCACTCGGGGCGCTACTCGGGCTGAGATCTCGATACCATCTC 307

QY 352 -----GGGGA 356
DB 308 TGGCCTTCCACAGTGGAGAGGGGCACCAAAAAGTAGGTCTTCTTACCCAGGGGA 367

QY 357 GACGGTAGTACAGCGGCTTCGAGGCAACAGCTGAGTGGGGGCGCAACACATGAT 416
DB 368 GACGGTAGTACAGCGGCTTCGAGGCAACAGCTGAGTGGGGGCGCAACACATGAT 427

QY 417 GGTGGAGTACGCGCGGGGCTCATCCCATCCCTGGGCTTCGCGCTGCCGACACTGT 476
DB 428 GGTGGAGTACGCGCGGGGCTCATCCCATCCCTGGGCTTCGCGCTGCCGACACTGT 487

QY 477 CTTGACGACCCAGGACTTCCTCACCCCTTCTATCTTCTGCTTCCTATGCTCGGGCCT 536

Db 488 CTTAGCACCAGGACTCTCTCACCTCTTCTATATCTCTTCTCTCTATGCTCGGGGCT 547
QY 537 CGGCTTGAGCTCACACCTACTCTTTGGCCAGGACCTTGA 579
Db 548 CCGGCTTGAGCTCACACCTACTCTTTGGCCAGGACCTTGA 590

RESULT 9
ADB58553
ID ADB58553 standard; DNA; 1590 BP.

AC ADB58553;

DT 04-DEC-2003 (first entry)

DE Toxicity-related gene, SEQ ID 3579.

XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.

XX Unidentified.

XX WO2003064624-A2.

XX 07-AUG-2003.

XX 31-JAN-2003; 2003WO-US003194.

XX 31-JAN-2002; 2002US-00060087.

XX 15-MAR-2002; 2002US-0364045P.

XX 15-MAR-2002; 2002US-0364055P.

XX 30-DEC-2002; 2002US-0436643P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX WPI; 2003-689530/65.

XX Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.

XX Claim 1; SEQ ID NO 3579; 1156pp; English.

XX The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1590 BP; 347 A; 435 C; 447 G; 361 T; 0 U; 0 Other;

XX Query Match 62.3%; Score 360.8; DB 10; Length 1590;

XX Best Local Similarity 75.4%; Pred. No. 2.1e-70; Indels 93; Gaps 1;

XX Matches 507; Conservative 0; Mismatches 72; Indels 93; Gaps 1;

QY 1 ATGCGATGGCGGTGGCGGGGGGTGGCGGGCGGCTGCTCTGCTGCTCGAGCG 60
Db 39 ATGCGGTGGCTGTGGCGGGGGGTGGCGATGATCACCCTGTTCTGACTATTGGCG 98

QY 61 GTGCTGACCCAGGTCGTCTGGCTCTGGGTACGACAGAGCTTCTGTTCCAGGCGAA 120
Db 99 GTGCTGATCCAGGCGGTCTGCTGTGGCTGGGTACTCAGAGCTTCTGTTCCAGAGANA 158
QY 121 GAGATAGCGCAGTTGGCGCGCAGTACGCTGGGCTGGACCAAGAGCTGGCTTCTCTGT 180
Db 159 GAGATAGCTCAGCTTGTCTGACAGTACGACGGGCTGGACCATGAGTGGCTTCTCTCG 218
QY 181 CTGATCTGTGAGCTGCGGCGGCTGACCCAGGCGACGCTGCTGCCGAGAGAGTGCAG 240
Db 219 CTGATCTGTGAGCTGCGGAGGCTGACCCAGGCGACGCTGCTGCCGAGTGCAG 278
QY 241 TGGGTGTTTGTGAATGCGGCTGGATGGCGGCTGATGGCGCATGTGCTTCTGACGCTCGCTG 300
Db 279 TGGGTGTTTGTGAACGCGGCGGCTGGATGGCGCATGTGCTTCTTTCAGCGCTCGCTG 338
QY 301 TCCGAGTATGCTGCTCTTTCGCGCACCGCTTGGGCTCCCGGGCCACTCG----- 351
Db 339 TCTGAGTACGTGCTCTTTCGCGCACCGCTTGGGCTCCCATGGCAATTCCGGGACGATAC 398
QY 352 ----- 351
Db 399 TGGCTGAGATTTTCAGATACCATCATCTCTGGCACTTTTCCACAGTGGAGAGAGGACCC 458
QY 352 -----GGGAGACGGTAGTACACGGGCTTGGTGGAGGCAACA 387
Db 459 ACAAAAAGTGAGTCTATTACCCAGGAGAGACGGTTGTCATGGACCTGGAGAGGCAACA 518
QY 388 GCTGTGAGTGGGGGCCAAACACATGGATGGTGGAGTACGGCCGGGGCGTCAATCCATCC 447
Db 519 GATGTGGAGTGGGACCAACACACGTGGATGTGGAGTATGGCCGGGGTGTATTCCGCT 578
QY 448 ACCCTGGCTTCGCGCTGGCGCGACACTGTCTTCAGCACCCAGGACTTCTCACCCCTCTTC 507
Db 579 ACCCTGGCATTTGCACCTTAGTGACACTATTTTCAGCACCCAGGACTTCTCACCCCTCTTC 638
QY 508 TATACTCTTCGCTCTCTATGCTCGGGGCTTCGGGCTTGAGCTCACCACCTACTCTTTGGC 567
Db 639 TATACCTTCGCGCTATGCTCGGGGCTTCGGGCTTGAGCTCACCACCTACTCTTTGGC 698
QY 568 CAGGACCCCTTGA 579
Db 699 CAAGACCCCTGA 710

RESULT 10
ACH50720
ID ACH50720 standard; cDNA; 442 BP.

XX ACH50720;

XX 13-OCT-2003 (first entry)

XX Human mammary gland cDNA #125.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 37932; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 442 BP; 102 A; 137 C; 111 G; 91 T; 0 U; 1 Other;
SQ
Query Match 27.4%; Score 158.4; DB 9; Length 442;
Best Local Similarity 99.4%; Pred. No. 1.2e-25;
Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 420 GGAGTACGGCGGGGGGTGATCCATCCACCTTGGCTTGGCGTGGCGACACTGTCCT 479
Db 1 GGAGTACGGCGGGGGGTGATCCATCCACCTTGGCGTGGCGGACACTGTCCT 60
QY 480 CAGCACCAGGACTTCTCTACCCCTTCTATCTCTTGGCTCTATGCTCGGGGCTCCG 539
Db 61 CAGCACCAGGACTTCTCTACCCCTTCTATCTCTTGGCTCTATGCTCTGTTGGGCTCCG 120
QY 540 GCTTGAGCTCACCACCTACCTCTTTGGCCAGGACCTTGA 579
Db 121 GCTTGAGCTCACCACCTACCTCTTTGGCCAGGACCTTGA 160
RESULT 11
ACA55628
ID ACA55628 standard; cDNA; 202 BP.
XX ACA55628;
XX
XX 06-JUN-2003 (first entry)
XX Human signalling pathway polynucleotide probe SEQ ID NO 226.
XX
XX Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
XX Homo sapiens.
XX
XX US6500938-B1.
XX
XX 31-DEC-2002.
XX
XX 30-JAN-1998; 98US-00016434.
XX
XX 30-JAN-1998; 98US-00016434.
PR

XX (INCY-) INCYTE GENOMICS INC.
XX Au-Young J, Seilhamer JJ;
XX WPI; 2003-352189/33.
XX
XX Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides.
XX
XX Claim 1; SEQ ID NO 226; 65pp; English.
XX
XX The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signalling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX
XX Sequence 202 BP; 22 A; 68 C; 76 G; 33 T; 0 U; 3 Other;
SQ
Query Match 21.9%; Score 127; DB 10; Length 202;
Best Local Similarity 98.4%; Pred. No. 9.7e-19;
Matches 127; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCAGTGGCGCGTGGCGCGGTGGCGGTGGCGCGCTGCTCTCTGCTGCTGCGAGCG 60
Db 74 ATGCAGTGGCGCGTGGCGCGGTGGCGGTGGCGCGCTGCTCTCTGCTGCTGCGAGCG 133
QY 61 GTGCTACCCAGGTCTGCTGGCTCTGGCTGGGTACCGAGAGCTTCGTTCCAGCGCGAA 120
Db 134 GTGCTGAGCCAGGTCTGCTGGCTCTGGCTGGGTACCGAGAGCTTCGTTCCAGCGCGAA 193
QY 121 GAGATAGCG 129
Db 194 GAGATAGCG 202
RESULT 12
ADI55424
ID ADI55424 standard; DNA; 202 BP.
XX ADI55424;
XX
XX 22-APR-2004 (first entry)
XX Human polynucleotide probe #226.
XX
XX Human; probe; ss; receptor-like polypeptide; transducing polypeptide;
KW effector-like polypeptide; cancer; immunopathology; neuropathology;
KW drug development; toxicology; carcinogenicity;
KW signalling pathway polypeptide; adrenal gland; bladder; bone;
KW bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;
KW diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;
KW dementia; amnesia; epilepsy; Alzheimer's disease; depression.
XX

OS Homo sapiens.
PN US2004010136-A1.
XX
PD 15-JAN-2004.
XX
PF 26-NOV-2002; 2002US-00305720.
XX
PR 30-JAN-1998; 98US-00016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
PI WPI; 2004-090520/09.
XX
DR
XX
XX New composition comprising polynucleotide probes, useful as array
PT elements in a microarray for monitoring the expression of target
PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic
PT fragments.
XX
PS Claim 6; SEQ ID NO 226; 73pp; English.
XX
XX The invention relates to a composition of polynucleotide probes
CC comprising first polynucleotide probes comprising at least a portion of a
CC gene encoding a receptor-like polypeptide, second polynucleotide probes
CC comprising at least a portion of a gene encoding a transducing
CC polypeptide and third polynucleotide probes comprising at least a portion
CC of a gene encoding an effector-like polypeptide. The probes of the
CC composition are useful as array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray is useful in the
CC diagnosis and treatment of cancer, an immunopathology or a
CC neuropathology. It can also be used for drug discovery and development,
CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.
CC Microarrays can also be used for monitoring the progression of diseases
CC that may be associated with the altered expression of signalling pathway
CC polypeptides. The composition can also be used to purify a subpopulation
CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile
CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of
CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,
CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or
CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,
CC epilepsy, Alzheimer's disease or depression. This sequence represents a
CC human polynucleotide probe of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 202 BP; 22 A; 68 C; 76 G; 33 T; 0 U; 3 Other;
Query Match 21.9%; Score 127; DB 12; Length 202;
Best Local Similarity 98.4%; Pred. No. 9.7e-19;
Matches 127; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCGATGGCGCGTGGCGCGGTGGCGCGGTGGCGCGGTCTCTGCTGCTGCAGCG 60
DB 74 ATGCGATGGCGCGTGGCGCGGTGGCGCGGTGGCGCGGTCTCTGCTGCTGCAGCG 133
QY 61 GTGCTGACCCAGGTCTGCTGGGTCTGGGTGAGTACGACAGAGTCTGCTCTTCCAGCGCGAA 120
DB 134 GTGCTGANCAGGTCTGCTGGGTCTGGGTGAGTACGACAGAGTCTGCTCTTCCAGCGCGNA 193
QY 121 GAGATAGCG 129
DB 194 GAGATAGCG 202
RESULT 13
ACH45312/c
ID ACH45312 standard; cDNA; 498 BP.
XX
AC ACH45312;
XX

DT 13-OCT-2003 (first entry)
XX Human foetal brain cDNA #6037.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 32524; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 498 BP; 108 A; 111 C; 157 G; 113 T; 0 U; 9 Other;
Query Match 16.1%; Score 93; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.7e-11;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 487 CAGGACTTCTCCACCCCTCTTCTATCTCTTCTGCTCTGCTCGGGGCTCGGGTTGAG 546
DB 498 CAGGACTTCTCCACCCCTCTTCTATCTCTTCTGCTCTGCTCGGGGCTCGGGTTGAG 439
QY 547 CTCACCACTTCTCTTTGGCCAGACCCCTTGA 579
DB 438 CTCACCACTTCTCTTTGGCCAGACCCCTTGA 406
RESULT 14
ADR02215
ID ADR02215 standard; DNA; 645 BP.
XX
AC ADR02215;
XX

DT 23-SEP-2004 (first entry)
XX A. gossypii genomic DNA PAG1566RP.
DE Filamentous funghi; ds; forensic identification; gene characterisation;
XX intergenomic comparison; chromosome mapping.
KW Eremothecium gossypii.
XX US6239264-B1.
XX 29-MAY-2001.
XX 24-DEC-1997; 97US-00998416.
XX 31-DEC-1996; 97CH-00000016.
XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX Philippe P, Pohlmann R, Steiner-Lange S, Mohr C, Wendland J;
PI Knechtle P, Reibschung C;
XX WPI; 1998-388120/33.
XX New gene for adenylate cyclase from *Ashbya gossypii* - useful for
PT generating recombinant microorganisms with alteration in gene of cAMP-
PT dependent signalling pathway for increasing production of fine chemicals.
XX Example 3; SEQ ID NO 907; 632pp; English.
XX The invention relates to isolated DNA molecules comprising isolated
CC genomic DNA sequences from the filamentous funghi *Ashbya gossypii*, the
CC sequences comprising ADR01309, ADR01366, ADR01367, ADR01386, ADR01428,
CC ADR01466, ADR01629, ADR01637, ADR02057, ADR02345 and ADR02369, chosen
CC from 1047 disclosed genomic sequences. Also included is a cloning vector
CC comprising a nucleotide sequence chosen from the above sequences. The
CC novel *Ashbya gossypii* genomic sequences are useful for forensic
CC identification, gene characterisation, for studying gene organisation by
CC intergenomic comparison (with *Saccharomyces cerevisiae*), identifying
CC biosynthetic genes for selectable markers, to isolate
CC promoters/terminators/centromeres, chromosome mapping, and in identifying
CC sequences unique to *Ashbya gossypii* for species identification. The
CC present sequence is an *A. gossypii* novel genomic sequence of the
XX invention.
XX Sequence 645 BP; 123 A; 196 C; 215 G; 111 T; 0 U; 0 Other;
SQ Query Match 11.3%; Score 65.4; DB 2; Length 645;
Best Local Similarity 68.7%; Pred. No. 5e-05; Indels 0; Gaps 0;
Matches 90; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 227 ACAGGAGTGCAGTGGGTGTCGTGAATGCGGTGGCTGGATGGGGCCATGCGCTTC 286
Db 106 ACCTGGAGGACGAGTGGGTGTTCAACACGGGGCGCGGTATGGGGCAGATGCTGATTC 165
QY 287 TGCACGCTCGCTGTCAGATGATGCTGCTTTTCGACACGCGCTTGGGCTCCGGGGCC 346
Db 166 TGCACGCGTCTGTGAGCGAGTACCTGATCTCTGTCGACGGCGGTGGGACGAGGGCC 225
QY 347 ACTCGGGGGAG 357
Db 226 ACACGGGGCGTG 236
RESULT 15
ADK65614
ID ADK65614 standard; DNA; 669 BP.
XX AC ADK65614;
XX 06-MAY-2004 (first entry)
XX Yeast C8-isomerase gene.

XX ds; gene; yeast; delta8-delta7-isomerase; enzyme; 7-dehydrocholesterol;
KW vitamin D3; skin cream; terpene; steroid hormone synthesis; C8-isomerase.
XX Saccharomyces cerevisiae.
XX Key Location/Qualifiers
FH 1. .669
FT /*tag= a
FT /product= "C8-isomerase"
XX DE10203352-A1.
XX 31-JUL-2003.
XX 29-JAN-2002; 2002DE-01003352.
XX 29-JAN-2002; 2002DE-01003352.
XX (BADI) BASF AG.
XX Lang C, Veen M;
XX WPI; 2003-732841/70.
DR P-PSDB; ADK65615.
XX Preparation of 7-dehydrocholesterol, useful as an intermediate for
PT vitamin D3, and its intermediates or products, comprises growth of
PT organisms with increased activity of specific enzymes.
XX Claim 15; Page 41-42; 120pp; German.
XX The present invention relates to a method for the preparation of 7-
CC dehydrocholesterol, and/or its biosynthetic intermediates and/or
CC secondary products, which comprises culturing organisms that have,
CC relative to the wild type, increased activity of at least one of delta8-
CC delta7-isomerase, delta5-desaturase and/or delta24-reductase. 7-
CC dehydrocholesterol is an intermediate for Vitamin D3, and its
CC intermediates/secondary products (e.g. zymosterol, lanosterol, squalene,
CC farnesol, geraniol, cholesta-trienol or tetraenol), and useful for
CC synthesis of terpenes, in dermatology and cosmetics, in synthesis of
CC saponins and steroid hormones, and as emulsifiers in skin creams. The
CC present sequence is the *S. cerevisiae* C8-isomerase gene as used in the
CC method of the invention.
XX Sequence 669 BP; 184 A; 151 C; 143 G; 191 T; 0 U; 0 Other;
SQ Query Match 8.9%; Score 51.4; DB 10; Length 669;
Best Local Similarity 65.0%; Pred. No. 0.064;
Matches 76; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 239 AGTGGGTGTTGCTGAATGCGGTGGCTGGATGGGGCCATGTCCTTCTGCACGCTCCG 298
Db 245 AATGGGTCTTCAACAATGCTGGTGGTGGATGGGCGCAATGATCATCTACAGCTTCCG 304
QY 299 TGTCCGAGTATGCTGCTCTTCGACACGCGCTTGGGCTCCCGGGGCACTCGGGG 355
Db 305 TATCCGAGTACTTAATTCCTGGAACCGCTGTTGGTACTGAAGGCGCACAGGTG 361
Search completed: October 30, 2005, 16:08:09
Job time : 487 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 15:25:50 ; Search time 166 Seconds
(without alignments)
5707.254 Million cell updates/sec

Title: US-09-823-069A-1
Perfect score: 579
Sequence: 1 atgcagtgccgcgtggccg.....tcttggccaggacccttga 579

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476	82.2	1700	2	US-08-928-612-2
2	127	21.9	202	4	US-09-016-434-226
3	65.4	11.3	645	3	US-08-998-416-907
4	51.4	8.9	689	4	US-09-614-221A-181
5	51.4	8.9	1508	1	US-08-240-496A-1
6	51.2	8.8	2700	4	US-09-902-540-7680
7	51.2	8.8	6351	4	US-09-902-540-757
8	50.4	8.7	4263	4	US-09-902-540-4659
9	50.4	8.7	29384	4	US-09-902-540-1229
10	49.6	8.6	1404	4	US-09-902-540-6859
11	49.6	8.6	3875	4	US-09-902-540-579
12	49	8.5	1016	4	US-09-949-016-2934
13	48.8	8.4	2556	4	US-09-902-540-4939
14	48.8	8.4	30780	4	US-09-902-540-1243
15	48.6	8.4	13335	4	US-09-949-016-14676
16	47.6	8.2	6645	4	US-09-902-540-8819
17	47.6	8.2	10210	4	US-09-902-540-938
18	47.4	8.2	636	4	US-09-692-945-3
19	46.8	8.1	2235	1	US-08-418-782-1
20	46.8	8.1	2235	2	US-08-228-662-1
21	46.8	8.1	2235	1	US-08-852-219-1
22	46.8	8.1	2331	1	US-08-418-782-20
23	46.8	8.1	2331	2	US-08-852-219-20
24	46.8	8.1	4403765	3	US-09-103-840A-2
25	46.8	8.1	4411529	3	US-09-103-840A-1
26	46	7.9	1185	4	US-09-252-991A-15968
27	46	7.9	1458	4	US-09-252-991A-16310

28	46	7.9	1599	4	US-09-252-991A-16536	Sequence 16536, A
29	45	7.8	9132	4	US-09-949-016-13070	Sequence 13070, A
30	44.6	7.7	77536	4	US-09-410-551B-1	Sequence 1, Appli
31	44.6	7.7	77536	4	US-09-940-316B-1	Sequence 1, Appli
32	44.4	7.7	34094	4	US-09-292-034-1	Sequence 1, Appli
33	44	7.6	4506	4	US-08-178-257-1	Sequence 1, Appli
34	44	7.6	6977	4	US-08-178-257-8	Sequence 8, Appli
35	43.8	7.6	2214	3	US-08-864-038A-1	Sequence 1, Appli
36	43.8	7.6	3331	3	US-08-864-038A-2	Sequence 2, Appli
37	43.8	7.6	3331	3	US-08-864-038A-4	Sequence 4, Appli
38	43.8	7.6	41106	4	US-09-949-016-15796	Sequence 15796, A
39	43.6	7.5	1410	4	US-09-902-540-3735	Sequence 3735, Ap
40	43.6	7.5	19112	4	US-09-902-540-1181	Sequence 1181, Ap
41	43.2	7.5	696	4	US-09-252-991A-647	Sequence 647, App
42	43.2	7.5	1371	4	US-09-252-991A-764	Sequence 764, App
43	43.2	7.5	1392	4	US-09-252-991A-704	Sequence 704, App
44	43.2	7.5	4826	4	US-09-772-304A-1	Sequence 1, Appli
45	43.2	7.5	10424	4	US-09-902-540-1015	Sequence 1015, Ap

ALIGNMENTS

RESULT 1
US-08-928-612-2
; Sequence 2, Application US/08928612
; Patent No. 5863766
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGMA RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,612
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0383 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTAT04
; CLONE: 1542751
US-08-928-612-2
Query Match 82.2%; Score 476; DB 2; Length 1700;
Best Local Similarity 86.2%; Pred. No. 4.8e-109;

CLASSIFICATION: 435
PRIOR APPLICATION DATA: CH 0016/97
APPLICATION NUMBER: 31-DEC-1996
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PE/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 907:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1566RP
US-08-998-416-907

Query Match 11.3%; Score 65.4; DB 3; Length 645;
Best Local Similarity 68.7%; Pred. No. 5.5e-07;
Matches 90; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 227 ACAGGAGTCAGTGGGTGTTGCTGAATGCGGGTGGCTGGATGGGGCCATGTGCCTTC 286
DB 106 ACCTGGAGGACGAGTGGGTGTTCAACAACGCGGGCGCGTATGGGGCAGATGCTGATTC 165
QY 287 TGCACGCTCGCTGCGAGTATGCTGCTCTTCGGCACCGCTTGGGCTCCCGGGCC 346
DB 166 TGCACGCGTGTGAGGAGTACTGCTGATTCCTGTCGACGCGCGCTGGGCGACGAGGGCC 225
QY 347 ACTCGGGGGAG 357
DB 226 ACACGGGCGTG 236

RESULT 4
US-09-614-221A-181
Sequence 181, Application US/09614221A
Patent No. 6723837
GENERAL INFORMATION:
APPLICANT: Karunanandaa, Balasulojini
APPLICANT: Yu, Jaehyuk
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
WITH STEROL SYNTHESIS AND METABOLISM
FILE REFERENCE: 16516.075
CURRENT APPLICATION NUMBER: US/09/614, 221A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/142,981
PRIOR FILING DATE: 1999-07-12
NUMBER OF SEQ ID NOS: 626
SEQ ID NO 181
LENGTH: 669
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-181

Query Match 8.9%; Score 51.4; DB 4; Length 669;
Best Local Similarity 65.0%; Pred. No. 0.0017;
Matches 76; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 239 AGTGGGTGTTTCGTAATGCGGGTGGCTGGATGGCGCATGTGCTTCTGCACGCTCGC 298
DB 245 AATGGGTCTTCAACAATGCTGGTGGTGGCGCAATGATCATCTACACGCTTCG 304
QY 299 TGTCCGAGTATGTGCTCTCTCGGCACCGCTTGGGCTCCCGGGCCACTCGGGGG 355
DB 305 TATCCGAGTACTTAATCTATTTCGGAACCGCTGTGTGTTGTTGTAAGGGGCACACAGGTG 361

RESULT 5
US-08-240-496A-1
Sequence 1, Application US/08240496A
Patent No. 5480805
GENERAL INFORMATION:
APPLICANT: Wolf, Fred R.
APPLICANT: Cuellar, Richard E.
TITLE OF INVENTION: Composition for Modulating Sterols in
TITLE OF INVENTION: Yeast
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Joanne M. Giesser
STREET: Amoco Corp., Suite 600, 55 Shuman Blvd.
CITY: Naperville
STATE: IL
COUNTRY: USA
ZIP: 60563
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,496A
FILING DATE: 10-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/929,764
FILING DATE: 12-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Giesser, Joanne M.
REGISTRATION NUMBER: 32,838
REFERENCE/DOCKET NUMBER: 31,456
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 717-2443
TELEFAX: (708) 717-2430
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 379..1047
US-08-240-496A-1

Query Match 8.9%; Score 51.4; DB 1; Length 1508;
Best Local Similarity 65.0%; Pred. No. 0.0021;
Matches 76; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 239 AGTGGGTGTTTCGTAATGCGGGTGGCTGGATGGCGCATGTGCTTCTGCACGCTCGC 298
DB 623 AATGGGTCTTCAACAATGCTGGTGGTGGCGCAATGATCATCTACACGCTTCG 682
QY 299 TGTCCGAGTATGTGCTCTCTCGGCACCGCTTGGGCTCCCGGGCCACTCGGGGG 355
DB 683 TATCCGAGTACTTAATCTATTTCGGAACCGCTGTGTGTTGTTGTAAGGGGCACACAGGTG 739

RESULT 6
US-09-902-540-7680
Sequence 7680, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 16:59:53 ; Search time 747 Seconds
(without alignments)

6401.547 Million cell updates/sec

Title: US-09-823-069A-1

Perfect score: 579

Sequence: 1 agcagtgccgctggccg.....tcttggccaggacccttga 579

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
- 25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	579	100.0	579	9	US-09-823-069-1
2	438.2	75.7	579	9	US-09-823-069-3
3	387	66.8	1535	9	US-09-925-298-37
4	387	66.8	1535	15	US-10-102-806-37
5	328	56.6	1569	22	US-10-764-420-2180

6	208.2	36.0	684	19	US-10-240-425-1015	Sequence 1015, Ap
7	158.4	27.4	442	10	US-09-918-995-37932	Sequence 37932, A
8	127	21.9	202	18	US-10-305-720-426	Sequence 226, App
9	93	16.1	498	10	US-09-918-995-32524	Sequence 32524, A
10	51.4	8.9	669	18	US-10-369-493-25589	Sequence 25589, A
11	51.4	8.9	669	20	US-10-793-639-181	Sequence 181, App
12	49.6	8.6	680	19	US-10-425-114-12614	Sequence 12614, A
13	49.6	8.6	1052	18	US-10-369-493-42833	Sequence 42833, A
14	49.6	8.6	1602	21	US-10-425-115-124288	Sequence 124288, A
15	49	8.5	636	20	US-10-384-339C-64	Sequence 64, Appl
16	49	8.5	1330	9	US-09-805-805-1	Sequence 1, Appli
17	49	8.5	1330	24	US-10-991-224-1	Sequence 1, Appli
18	48.8	8.4	1632	19	US-10-425-114-24893	Sequence 24893, A
19	48.8	8.4	1636	19	US-10-425-114-2037	Sequence 2037, Ap
20	48.8	8.4	1693	21	US-10-425-115-78945	Sequence 78945, A
21	47.4	8.2	633	9	US-09-817-814-1	Sequence 1, Appli
22	47.4	8.2	633	17	US-10-011-364-1	Sequence 1, Appli
23	47.4	8.2	633	17	US-10-174-394-1	Sequence 1, Appli
24	47.4	8.2	633	17	US-10-174-394-47	Sequence 47, Appl
25	47.4	8.2	633	17	US-10-174-394-49	Sequence 49, Appl
26	47.4	8.2	633	17	US-10-174-394-51	Sequence 51, Appl
27	47.4	8.2	633	17	US-10-174-394-53	Sequence 53, Appl
28	47.4	8.2	633	18	US-10-321-962-1	Sequence 1, Appli
29	47.4	8.2	633	21	US-10-435-087-3	Sequence 3, Appli
30	47.4	8.2	633	21	US-10-435-087-35	Sequence 35, Appl
31	47.4	8.2	633	22	US-10-948-588-1	Sequence 1, Appli
32	47.4	8.2	633	24	US-10-980-764-1	Sequence 1, Appli
33	47.4	8.2	633	24	US-10-980-764-21	Sequence 21, Appl
34	47.4	8.2	633	24	US-10-980-659-1	Sequence 1, Appli
35	47.4	8.2	633	24	US-10-980-659-21	Sequence 21, Appl
36	47.4	8.2	633	24	US-10-980-695-1	Sequence 21, Appl
37	47.4	8.2	633	24	US-10-980-695-21	Sequence 1, Appli
38	47.4	8.2	633	26	US-11-027-948-1	Sequence 1, Appli
39	47.4	8.2	636	14	US-10-005-646-1	Sequence 4, Appli
40	47.4	8.2	636	14	US-10-090-983-4	Sequence 4, Appli
41	47.4	8.2	636	17	US-10-174-394-33	Sequence 33, Appl
42	47.4	8.2	636	21	US-10-435-087-21	Sequence 21, Appl
43	47.4	8.2	636	24	US-10-980-764-5	Sequence 5, Appli
44	47.4	8.2	636	24	US-10-980-659-5	Sequence 5, Appli
45	47.4	8.2	636	24	US-10-980-695-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-823-069-1
; Sequence 1, Application US/09823069
; Patent No. US20020061847A1
; GENERAL INFORMATION:
; APPLICANT: Wheeler, Kenneth
; APPLICANT: Mach, Robert
; APPLICANT: Childers, Steven
; APPLICANT: Shelness, Gregory
; APPLICANT: Wang, Li-Ming
; TITLE OF INVENTION: METHODS AND COMPOSITIONS UTILIZING AN ALTERNATIVE SPLICE VARIANT
; TITLE OF INVENTION: SIGMA-1 RECEPTOR
; FILE REFERENCE: 9151.6
; CURRENT APPLICATION NUMBER: US/09/823,069
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,694
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(579)
US-09-823-069-1

Qy 480 CAGCACCAGGACTTCCTCACCCCTTCTATACCTTGCTCCTATGCTCGGGGCTCCG 539
|||||
Dβ 61 CAGCACCAGGACTTCCTCACCCCTTCTATACCTTGCTCCTATGCTCGGGGCTCCG 120

QY 540 GCTTGAGCTCACCACCTACCTCTTTGGCCAGGACCTTGA 579
 121 GCTTGAGCTCACCACCTACCTCTTTGGCCAGGACCTTGA 160

```

RESULT 8
US-10-305-720-226
; Sequence 226, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 226
; LENGTH: 202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 1542751
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (202)
; OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-226

```

	Query Match	21.9%;	Score 127;	DB 18;	Length 202;
	Best Local Similarity	98.4%;	Pred. No. 2e-24;		
	Matches 127;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	ATGCAGTGGCCGCTGGGCGCGCTGGCGCTGCTCTGGCTGTGTCGACGCG	60		
Db	74	ATGCAGTGGCCGCTGGGCGCGCTGGCGCTGCTCTGGCTGTGTCGACGCG	133		
Qy	61	GTGCTGACCCAGGTCGTCTGGCTTGGCTGGGTTACGCAGAGCTTCTGTTCCAGCGGAA	120		
Db	134	GTGCTGANCAGGTCGTCTGGCTTGGCTGGGTTACGCAGAGCTTCTGTTCCAGCGGAA	193		
Qy	121	GAGATAGCG	129		
Db	194	GAGATAGCG	202		

```

RESULT 9
US-09-918-995-32524/c
; Sequence 32524, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32524
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (1)...(498)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32524

Query Match      16.1%; Score 93; DB 10; Length 498;
Best Local Similarity 100.0%; Pred. No. 3e-15;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      487 CAGGACTTCCTCACCCCTCTTCTATACACTTCGCTTCCTATGCTCGGGGCTCCGGCTTGAG 546
Db      498 CAGGACTTCCTCACCCCTCTTCTATACACTTCGCTTCCTATGCTCGGGGCTCCGGCTTGAG 439

Qy      547 CTCACCACTACCTCTTTGGCCAGGACCCCTTGA 579
Db      438 CTCACCACTACCTCTTTGGCCAGGACCCCTTGA 406

RESULT 10
US-10-369-493-25589
; Sequence 25589, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25589
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-25589

```

```

RESULT 11
US-10-793-639-181
; Sequence 181, Application US/10793639
; Publication No. US20040199940A1
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; WITH STEROL SYNTHESIS AND METABOLISM
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/10/793,639
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US/09/614,221A
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626

```

; SEQ ID NO 181
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-793-639-181

Query Match 8.9%; Score 51.4; DB 20; Length 669;
Best Local Similarity 65.0%; Pred. No. 0.0005;
Matches 76; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 239 AGTGGGTGTTCTGTAATGCGGGTGGCTGTGATGGGCGCCATGTCCTTCTGCACGCTCGC 298
DB 245 AATGGTCTTCAACATGCTGGTGGCGATGGCCAAATGATCATCTTACACGCTTCGG 304
QY 299 TGTCCGAGTATGCTGCTTCTCGGCACCGCTTGGGCTCCGCGGCCACTCGGGG 355
DB 305 TATCCGAGTACTTAATTTCTATTCGGAACCGCTGTTGGTACTGAAGGCGCACAGGTG 361

RESULT 12

US-10-425-114-12614
; Sequence 12614, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12614
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Zea mays

; FEATURE:
; OTHER INFORMATION: Clone ID: 701185825_FLI
US-10-425-114-12614

Query Match 8.6%; Score 49.6; DB 19; Length 680;
Best Local Similarity 45.9%; Pred. No. 0.0015;
Matches 169; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 6 GTGGGCGTGGGCGGCGGTGGGCGGCTGCTCCCTGGCTGTCGACGCGTGCT 65
DB 40 GTGCTGGGGCGGTGCTCCGCGCGCACGCGGTGGCGGTGCTGCTCGCTCGTTCCA 99
QY 66 GACCCAGGTGCTGCTGGCTGGGTACGACAGCTTCTGTTCCAGCGCGAAGAGAT 125
DB 100 GCTCTGTTATGCTGCTCGGNAGCCCTTGTCAGAGAGCGGTGCAGCTGTGGAGAC 159
QY 126 AGCGCATGTTGGCGCGCAGTACGCTGGGTGGACACAGAGTGGCTTCTCTGCTGAT 185
DB 160 CTTGTCGTTGGCGAGGTGGCCGTTTTCGCGCGTGCCTGGCGCTCATCGACAGGAC 219
QY 186 CTGGAGCTGGCGGCTGCACCCAGGACCGTGTGCGCGGAGGAGCTCAGTGGGT 245
DB 220 GGATCGAGCGGGGTGCTGCCCCACGCGAGGCGCGCGGTGGGCTGGCGATGCT 279
QY 246 GTTCTGTAATGCGGGTGGCTGGAATGGCGCCATGCTCTTCTGACGCTCGCTGCCA 305
DB 280 GGGCGCTTCTGCTGGGTTCGCGCGGCGAGGTGTGCACAGTGGAAACGCTGTACCG 339
QY 306 GTATGCTGCTTCTCGGCACCGCTTGGGCTCCCGCGGCCACTCGGGGGAGACGGTAGT 365
DB 340 GCAGGCGCGCTGCTGAGGCGCGACCGGAGCTCTTCTTGGACGGCTCGGCGGCGG 399
QY 366 ACACGGGC 373

DB 400 CTTGGGCC 407

RESULT 13

US-10-369-493-42833
; Sequence 42833, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 42833
; LENGTH: 1052
; TYPE: DNA

; ORGANISM: Myxococcus xanthus
US-10-369-493-42833

Query Match 8.6%; Score 49.6; DB 18; Length 1052;
Best Local Similarity 48.3%; Pred. No. 0.0015;
Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 46 CTGGCTGTGCGAGCGGTGCTGACCCAGCTGCTGTGGCTGCTGCTGGTACGACGAGCTTC 105
DB 628 CTCACGGCGAGGTGCCGCTGGGACGTTTCGACCCCGCTCGCGCGCAAGCGGGCTG 687
QY 106 GTCTTCCAGCGCAAGAGATAGCGAGTTGGCGCGGAGTAGCTGGGCTGGACCAACGAG 165
DB 688 GACTCGCGGTGAGCGCCATGTCACGCGGTGCTCAAGCCGAGCCCGAGGACCGCTAC 747
QY 166 CTGGCCTTCTCTGCTGATGCTGAGCTGCGGGCGGCTGCACCCAGGCGACGCTGCTGCC 225
DB 748 CCTCCGTCACCACTCATCGCGACCTTGGAGCTGCTGCTCCGCGGACGCTGCGGTCA 807
QY 226 GACGAGGAGCTCAGTGGGTGTTCTGTAATGCGGTGGTGGATGCGGCGGCTGCGCTT 285
DB 808 CTGGCGCGGTGAAGCTACGACCTCCAGCGCGTCCGCGGCTGTCGCAAGGTGGTG 867
QY 286 CTGCACGCTCGCTGTCGAGTATGCTGCTCTTTCGCGCACCGCTTG 333
DB 868 CGCATGGCCCTGCAGGCGCGGCGGTGCTGCTGCTGGTGGCGGCGCTG 915

RESULT 14

US-10-425-115-124288
; Sequence 124288, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 124288
; LENGTH: 1602
; TYPE: DNA

; ORGANISM: Zea mays
; FEATURE:

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 30, 2005, 13:54:41 : Search time 2871 Seconds
(without alignments)
7676.493 Million cell updates/sec

Title: US-09-823-069a-1
Perfect score: 579
Sequence: 1 atgcagtgccgcgtggccg.....tctttggccagacccttga 579

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	579	100.0	852	4	BI553398 603193344
2	483	83.4	1022	5	BQ070845 AGENCOURT
3	483	83.4	1077	5	BX416039 BX416039
4	483	83.4	1608	3	CR606713 full-leng
5	481.4	83.1	890	5	BX463121 BX463121
6	478.8	82.7	900	4	EG324967 602423618
7	478.2	82.6	850	4	BI196388 602756438
8	476	82.2	835	5	BQ230931 AGENCOURT
9	476	82.2	869	1	AL521606 AL521606
10	476	82.2	884	1	AL543500 AL543500
11	476	82.2	894	1	AL532097 AL532097
12	476	82.2	895	5	BQ954246 AGENCOURT
13	476	82.2	923	5	BU177987 AGENCOURT
14	476	82.2	941	5	BQ899240 AGENCOURT
15	476	82.2	957	5	BX418964 BX418964
16	476	82.2	1003	1	AL537799 AL537799
17	476	82.2	1005	5	BU190210 AGENCOURT
18	476	82.2	1018	5	BX376684 BX376684
19	476	82.2	1024	5	BX462186 BX462186
20	476	82.2	1046	3	CR622156 full-leng
21	476	82.2	1073	1	AL583269 AL583269
22	476	82.2	1104	5	BX427615 BX427615
23	476	82.2	1129	1	AL528279 AL528279
24	476	82.2	1538	3	CR604206 full-leng

25	476	82.2	1560	3	CR594423
26	476	82.2	1560	3	CR608230 full-leng
27	476	82.2	1589	3	CR606455 full-leng
28	476	82.2	1592	3	CR604448 full-leng
29	476	82.2	1601	3	CR596218 full-leng
30	476	82.2	1606	3	CR609235 full-leng
31	476	82.2	1618	3	CR608948 full-leng
32	476	82.2	1619	3	CR625524 full-leng
33	476	82.2	1627	3	CR590434 full-leng
34	476	82.2	1627	3	CR602297 full-leng
35	476	82.2	1629	3	CR595153 full-leng
36	476	82.2	1631	3	CR611273 full-leng
37	476	82.2	1640	3	CR619527 full-leng
38	476	82.2	1644	3	CR603042 full-leng
39	476	82.2	1646	3	CR592762 full-leng
40	476	82.2	1654	3	CR611989 full-leng
41	476	82.2	1658	3	CR598930 full-leng
42	476	82.2	1659	3	CR593849 full-leng
43	476	82.2	1671	3	CR592515 full-leng
44	475.6	82.1	1093	1	AL520391 AL520391
45	474.4	81.9	916	1	AL531582 AL531582

ALIGNMENTS

RESULT 1
BI553398
LOCUS 603193344F1 NTH_MGC_95 Homo sapiens cDNA clone IMAGE:5264674 5',
DEFINITION mRNA sequence.
ACCESSION BI553398
VERSION BI553398.1 GI:15440710
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 852)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: L1AM11667 row: g column: 11
High quality sequence stop: 824.

FEATURES
source

1. .852
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5264674"
/tissue type="hippocampus"
/lab_host="DH10B"
/clone_l1b="NIH_MGC_95"
/notes="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

DEFINITION BX416039 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CS0DA001Y009 5-PRIME, mRNA sequence.

ACCESSION BX416039

VERSION BX416039.2 GI:46954005

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1077)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30763590.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5104.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DA001AH05QP1&c=5104.r.

FEATURES
source
1..1077
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA001Y009"
/issue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 83.4%; Score 483; DB 5; Length 1077;
Best Local Similarity 88.6%; Pred. No. 2.6e-100;
Matches 554; Conservative 0; Mismatches 25; Indels 46; Gaps 1;
1 ATGCAGTGGGCCCTGGGCGCGGCTGGGCGCGCTGCTCTCTGGCTGTCGAGCG 60
2 ATGCAGTGGGCCCTGGGCGCGGCTGGGCGCGGCTGCTCTCTGGCTGTCGAGCG 141
61 GTGCTGACCCAGCTGCTGGCTCTGGCTGGGTACGACGAGCTTCTTCCAGCGCGAA 120
142 GTGCTGACCCAGCTGCTGGCTCTGGCTGGGTACGACGAGCTTCTTCCAGCGCGAA 201
121 GAGATAGCGAGTGGCGCGGCTGACGCTGGGCTGGACACGAGCTGGGCTTCTCTCGT 180
202 GAGATAGCGAGTGGCGCGGCTGACGCTGGGCTGGACACGAGCTGGGCTTCTCTCGT 261
181 CTGATCGTGAAGTGGCGCGGCTGACCCAGGCGCAGTGTGCTGCCGACGAGGAGCTGCAG 240
262 CTGATCGTGAAGTGGCGCGGCTGACCCAGGCGCAGTGTGCTGCCGACGAGGAGCTGCAG 321
241 TGGGTGTTCTGAATGCGGGTGGCTGATGGGCGGCATGTGCTTCTGACGCGCTCGCTG 300
322 TGGGTGTTCTGAATGCGGGTGGCTGATGGGCGGCATGTGCTTCTGACGCGCTCGCTG 381
301 TCCGAGTATCTGCTGCTCT-----319
382 TCCGAGCGGCTACTGGGCTGATCTCGGATACCATCATCTCTGGCACCTTCCACCATGTG 441
320 -----TCGGCACCGCTTGGGCTCCCGCGCCACTCGGGGAGACGGTAGTACACGGGCC 374
442 GAGAGAGGGGCACCAACAAAGTGAGGTCTTTACCCAGGGGAGACGGTAGTACACGGGCC 501

QY 375 TGGTGGAGCAACAGCTGTGGAGTGGGGGCGCAACACATGGATGGTGGAGTACGGCGGG 434
DB 502 TGGTGGAGCAACAGCTGTGGAGTGGGGGCGCAACACATGGATGGTGGAGTACGGCGGG 561
QY 435 CGTCATCCCATCCACCCCTTGGCCTTCGCGCTGGCCGACACTGCTTTCAGCACCCAGGACTT 494
DB 562 CGTCATCCCATCCACCCCTTGGCCTTCGCGCTGGCCGACACTGCTTTCAGCACCCAGGACTT 621
QY 495 CTTCAACCCCTTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 554
DB 622 CTTCAACCCCTTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 681
QY 555 CTACCTCTTTGGCCAGGACCCCTTGA 579
DB 682 CTACCTCTTTGGCCAGGACCCCTTGA 706

RESULT 4
CR606713 1608 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DA001Y009 of Neuroblastoma of Homo
DEFINITION sapiens (human).
ACCESSION CR606713
VERSION CR606713.1 GI:50487520
KEYWORDS HTC; CNSLT.CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1608)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1608)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1608
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA001Y009"
/issue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 83.4%; Score 483; DB 3; Length 1608;
Best Local Similarity 88.6%; Pred. No. 2.7e-100;
Matches 554; Conservative 0; Mismatches 25; Indels 46; Gaps 1;
1 ATGCAGTGGGCCCTGGGCGCGGCTGGGCGCGGCTGGGCGCGGCTGCTCTCTGGCTGTCGAGCG 60
DB 82 ATGCAGTGGGCCCTGGGCGCGGCTGGGCGCGGCTGGGCGCGGCTGCTCTCTGGCTGTCGAGCG 141
QY 61 GTGCTGACCCAGCTGCTGGCTCTGGCTGGGTACGACGAGCTTCTTCCAGCGCGAA 120
DB 142 GTGCTGACCCAGCTGCTGGCTCTGGCTGGGTACGACGAGCTTCTTCCAGCGCGAA 201
QY 121 GAGATAGCGAGTGGCGCGGCTGACGCTGGGCTGGACACGAGCTGGGCTTCTCTCGT 180
DB 202 GAGATAGCGAGTGGCGCGGCTGACGCTGGGCTGGACACGAGCTGGGCTTCTCTCGT 261

```
QY 181 CTGATCGTGGAGCTCGCGGGCTGCACCCAGGCCACGTGCTGCCGACGAGGAGCTGCAG 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
262 CTGATCGTGGAGCTCGCGGGCTGCACCCAGGCCACGTGCTGCCGACGAGGAGCTGCAG 321
QY 241 TGGGTGTTTCGTGAATCGCGGTGGCTGGATGGCGGCCATGTGCTTCTTCAGCGCCTCGCTG 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
322 TGGGTGTTTCGTGAATCGCGGTGGCTGGATGGCGGCCATGTGCTTCTTCAGCGCCTCGCTG 381
QY 301 TCCGAGTATGCTGCTCT----- 319
Db |||||||
382 TCCGAGGGGCTACTCGGGCTGAGATCTCGGATACCATCATCTCTGCGACCTTCCACAGTG 441
QY 320 -----TCGCGACCGCTTGGGCTCCGCGGCCACTCGGGGAGAGCGTAGTACAGGGCC 374
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
442 GAGAGAGGCGACCAACAAAGTGAAGTCTTTACCCAGGGGAGACGGTAGTACAGGGCC 501
QY 375 TGGTGAGGCAACAGCTGTGGAGTGGGGGCCAAACACATGGATGGTGGAGTACGGCGGGG 434
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
502 TGGTGAGGCAACAGCTGTGGAGTGGGGGCCAAACACATGGATGGTGGAGTACGGCGGGG 561
QY 435 CGTCAATCCCATCCACCTGGCCTTCGCGTGGCGGACACATGCTTTCAGGACCCAGGACTT 494
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
562 CGTCAATCCCATCCACCTGGCCTTCGCGTGGCGGACACATGCTTTCAGGACCCAGGACTT 621
QY 495 CCTCACCTCTCTATACCTTCGCTCCCTATGCTCGGGGCCCTCCGGCTTGAGCTCACCAC 554
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
622 CTCACCTCTCTCTATACCTTCGCTCCCTATGCTCGGGGCCCTCCGGCTTGAGCTCACCAC 681
QY 555 CTACCTCTTTGGCCAGGACCCCTTGA 579
Db |||||||
682 CTACCTCTTTGGCCAGGACCCCTTGA 706
```

```
RESULT 5
BX463121
LOCUS BX463121 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM007Y109 5-PRIME, mRNA sequence.
ACCESSION BX463121
VERSION BX463121.2 GI:47064026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 890)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31027522.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5104.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0DM007AB05QPI&c=5104.r.
```

```
FEATURES
Location/Qualifiers
1. .890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007Y109"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
```

enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

```
Query Match 83.1%; Score 481.4; DB 5; Length 890;
Best Local Similarity 88.5%; Pred. No. 6e-100;
Matches 553; Conservative 0; Mismatches 26; Indels 46; Gaps 1;

QY 1 ATGCAGTGGGCGGTGGGCGCGGTGGGCGGTGGGCGCGGCTGCTCTCGCTGTTCGAGGG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
69 ATGCCGTGGGCGGTGGGCGCGGTGGGCGCGGCTGCTCTGGCTGTTCGAGCG 128
QY 61 GTGCTGACCCAGGTGCTGCTGGCTGCTGGCTGGGTAGCAGAGCTTCTCTCCAGCGCAA 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
129 GTGCTGACCCAGGTGCTGCTGGCTGCTGGCTGGGTAGCAGAGCTTCTCTTCAGCGCGNA 188
QY 121 GAGATAGCGCATGTGGGCGCGGACGTACGCTGGGCTGGACCAAGAGCTGGCCTTCTCTGT 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
189 GAGATAGCGCATGTGGGCGCGGACGTACGCTGGGCTGGACCAAGAGCTGGCCTTCTCTGT 248
QY 181 CTGATCGTGGAGCTGGGCGCGGCTGACCCAGGCGACGTGCTGCCGACGAGAGCTGCAG 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
249 CTGATCGTGGAGCTGGGCGCGGCTGACCCAGGCGACGTGCTGCCGACGAGAGCTGCAG 308
QY 241 TGGCTGTTTCGTGAATGGGGTGGCTGGATGGGCGGCGCATGCTTCTCGAGCCTCGCTG 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
309 TGGGTGTTTCGTGAATGGGGTGGCTGGATGGGCGGCGCATGCTTCTCGAGCCTCGCTG 368
QY 301 TCCGAGTATGCTGCTCT----- 319
Db |||||||
369 TCCGAGGCGCTACTGGGCTGAGATCTCGGATACCATCATCTCTGCGACCTTCCACAGTG 428
QY 320 -----TCGCGACCGCTTGGGCTCCCGGGCGCATCTCGGGGGAGAGCGGTAGTACAGGGCC 374
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
429 GAGAGAGGCGCACCAACAAAGTGAGGTCTTCTACCCAGGGGAGACGGTAGTACAGGGCC 488
QY 375 TGGTGAGGCAACAGCTGTGGAGTGGGGGCCAAACACATGGATGGTGGAGTACGGCGGGG 434
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
489 TGGTGAGGCAACAGCTGTGGAGTGGGGGCCAAACACATGGATGGTGGAGTACGGCGGGG 548
QY 435 CGTCAATCCCATCCACCTGGGCTTTCGCGCTGGCGGACACATGCTTCTTCAGCACCAGGACTT 494
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
549 CGTCAATCCCATCCACCTGGGCTTTCGCGCTGGCGGACACATGCTTTCAGCACCAGGACTT 608
QY 495 CCTCACCTCTCTATACCTTCTGCTCTCTATGCTCGGGGCGCTTCGGCTTGAGCTCACCAC 554
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
609 CCTCACCTCTCTATACCTTCTGCTCTCTATGCTCGGGGCGCTTCGGGCTTGAGCTCACCAC 668
QY 555 CTACCTCTTTGGCCAGGACCCCTTGA 579
Db |||||||
669 CTACCTCTTTGGCCAGGACCCCTTGA 693
```

RESULT 6

```
BX324967
LOCUS 602423618F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4561493 5',
DEFINITION mRNA sequence.
ACCESSION BG324967
VERSION BG324967.1 GI:13131508
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 900)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
```



```
Db      585 GCTGTGGAGTGGGGCCAAACACATGATGTGTGAGTAGCGGCGGGCGTTCATCCATCC 644
Qy      448 ACCCTGGGCTTCGGCGTGGCCGACACTGTCTTCAGACCCAGGACTTCTCTACCCCTCTTC 507
Db      645 ACCCTGGGCTTCGGCGTGGCCGACACTGTCTTCAGACCCAGGACTTCTCTACCCCTCTTC 704
Qy      508 TATACCTCTCGCTCTATGCTCGGGGCTTCGGGCTTGAGCTCACCACTACTCTTTGGC 567
Db      705 TATACCTCTCGCTCTATGCTCGGGGCTTCGGGCTTGAGCTCACCACTACTCTTTGGC 764
Qy      568 CAGGACCCCTTGA 579
Db      765 CAGGACCCCTTGA 776

RESULT 11
AL532097
LOCUS   AL532097 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION
ACCESSION
VERSION AL532097.3 GI:45707026
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 894)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:31069929.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5104.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?B=CS0DM004DF02QPI&c=5104.r.
Location/Qualifiers
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004Y104"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

FEATURES
source
ORIGIN
Query Match 82.2%; Score 476; DB 1; Length 894;
Best Local Similarity 86.2%; Pred. No. 1e-98;
Matches 579; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

Qy      1 ATGCAGTGGCGGTGGCGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGT 60
Db      77 ATGCAGTGGCGGTGGCGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGT 136
Qy      61 GTGCTGACCCAGGTGCTGTGGCTCTGGCTGGGTAGCAGAGCTTGTCTTCCAGCGCGAA 120
Db      137 GTGCTGACCCAGGTGCTGTGGCTCTGGCTGGGTAGCAGAGCTTGTCTTCCAGCGCGAA 196
Qy      121 GAGATAGCGCAATTGGCGCGGACAGTAGCGTGGGCTGGACCAAGAGCTGGCCCTTCTCTCGT 180
```

```
Db      197 GAGATAGCGCAATTGGCGCGGACAGTACGCTGGGCTGGACCAAGAGCTGGCCCTTCTCTCGT 256
Qy      181 CTGATCGTGGAGCTGCGCGCGCTGCACCCAGGCGACAGTGTCTGCCGACGAGAGAGTGCAG 240
Db      257 CTGATCGTGGAGCTGCGCGCGCTGCACCCAGGCGACAGTGTCTGCCGACGAGAGAGTGCAG 316
Qy      241 TGGGTGTTCTGTAATGCGGGTGGCTGGATGGGCGGCATGTGCTTCTTCACAGCCTCGCTG 300
Db      317 TGGGTGTTCTGTAATGCGGGTGGCTGGATGGGCGGCATGTGCTTCTTCACAGCCTCGCTG 376
Qy      301 TCCGAGTATGTGCTGCTCTTCGGCACCGCTTGGGCTCCCGCGGCACTCG ----- 351
Db      377 TCCGAGTATGTGCTGCTCTTCGGCACCGCTTGGGCTCCCGCGGCACTCGGGGGCGCTAC 436
Qy      352 ----- 351
Db      437 TGGGCTGAGATCTCGGATACCATCTCTTGGCACCTTCCACAGTGGAGAGAGGGCACCC 496
Qy      352 -----GGGAGAGCGGTAGTACACGGGCGCTGGTGAGGCAACA 387
Db      497 ACCAAAAGTGAGGTCTTCTACCCAGGGGAGACGGTAGTACACGGGCGCTGGTGAGGCAACA 556
Qy      388 GCTGTGGAGTGGGGCCAAACACATGGATGGTGGAGTACGGCGGGGGCGTTCATCCATCC 447
Db      557 GCTGTGGAGTGGGGCCAAACACATGGATGGTGGAGTACGGCGGGGGCGTTCATCCATCC 616
Qy      448 ACCCTGGCCTTTCGGCGCTGGCGGACACTGTCTTACGACACCCAGGACTTCTCTACCCCTTTC 507
Db      617 ACCCTGGCCTTTCGGCGCTGGCGGACACTGTCTTACGACACCCAGGACTTCTCTACCCCTTTC 676
Qy      508 TATACCTCTTCGCTCTATGCTCGGGGCGCTCGGCTTGAGCTCACCACTACTCTTGGC 567
Db      677 TATACCTCTTCGCTCTATGCTCGGGGCGCTCGGCTTGAGCTCACCACTACTCTTGGC 736
Qy      568 CAGGACCCCTTGA 579
Db      737 CAGGACCCCTTGA 748

RESULT 12
BQ954246 895 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8794813 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374677
DEFINITION
ACCESSION BQ954246
VERSION BQ954246.1 GI:22369724
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 895)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2554 row: a column: 14
High quality sequence stop: 665.
Location/Qualifiers
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6374677"
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 04:28:08 ; Search time 167 Seconds
(without alignments)
444.659 Million cell updates/sec

Title: US-09-823-069a-2
Perfect score: 1016
Sequence: 1 MQWVGRRWAAALLAVAA.....RSVARGRLRLTYLFGQDP 192

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1016	100.0	192	4	Aau09154 Human sig
2	990.5	97.5	223	2	Aay03755 Human sig
3	990.5	97.5	223	3	Aab19077 Amino aci
4	990.5	97.5	223	7	Ades4607 Human pro
5	990.5	97.5	223	8	Abm81032 Tumour-as
6	934.5	92.0	223	7	Ades4605 Rat Prote
7	925	91.0	192	4	Aau09155 Mouse sig
8	827.5	81.4	195	3	Aab58747 Breast an
9	544	53.5	106	5	Abb06121 Human NS
10	212.5	20.9	222	2	Aar90670 S. cerevi
11	212.5	20.9	222	7	Adk65615 Yeast C8-
12	212.5	20.9	222	8	Adn19249 Bacterial
13	206	20.3	219	8	Adn19490 Bacterial
14	204	20.1	285	8	Adn21450 Bacterial
15	180	17.7	259	8	Ads24370 Bacterial
16	91.5	9.0	395	8	Adq96108 T cell ac
17	91.5	9.0	395	8	Adq96216 T cell ac
18	88	8.7	442	4	Aab93553 Human pro
19	88	8.7	442	8	Abm81907 Tumour-as
20	86.5	8.5	414	4	Aam93876 Human pol
21	86.5	8.5	414	4	Adl31958 Human pro
22	86	8.5	471	6	Abp57482 Mycobacte
23	83	8.2	194	2	Aar47338 Peptide f
24	83	8.2	1157	5	Aam51707 Mouse TRP
25	83	8.2	1158	4	Aab86166 Mouse MTR

26	83	8.2	1158	4	AAB86164	Aab86164 Mouse MTR
27	83	8.2	1158	5	ABR83854	Abbr83854 Mouse ltr
28	83	8.2	1158	8	ADR87163	Adr87163 Mouse bet
29	82	8.1	1045	7	ABO68915	AbO68915 Pseudomon
30	81.5	8.0	440	6	ABU34006	Abu34006 Protein e
31	81.5	8.0	441	6	ADA34139	Ada34139 Acinetoba
32	81.5	8.0	487	6	ADA35480	Ada35480 Acinetoba
33	81.5	8.0	523	8	ADS23855	Ads23855 Bacterial
34	81	8.0	442	4	ABR95029	AbR95029 Human pro
35	81	8.0	463	7	ABO70635	AbO70635 Pseudomon
36	81	8.0	630	4	AAB83244	Aab83244 Human FAT
37	80.5	7.9	311	3	AAG05169	Aag05169 Arabidops
38	80.5	7.9	323	3	AAG05168	Aag05168 Arabidops
39	80.5	7.9	353	4	AAE09319	Aae09319 Soybean a
40	80.5	7.9	494	7	ABO83481	AbO83481 Pseudomon
41	80.5	7.9	861	7	ADM26091	Adm26091 Hyperther
42	80	7.9	1586	7	ADM26435	Adm26435 Hyperther
43	79.5	7.8	491	6	ABU50481	Abu50481 Protein e
44	79	7.8	330	4	ABG67521	Abg67521 Amino aci
45	78.5	7.7	469	8	ADS27918	Ads27918 Bacterial

ALIGNMENTS

RESULT 1
AAU09154
ID AAU09154 standard; protein; 192 AA.
XX
AC AAU09154;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human sigma 1 receptor splice variant, sigma1beta.
XX
XX Human; sigma 1 receptor; sigma1beta; tumour imaging; cancer;
KW cell proliferation disorder; tumour; diagnostic; cycostatic.
XX
OS Homo sapiens.
XX
PN WO200174297-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US010650.
XX
PR 31-MAR-2000; 2000US-0193694P.
XX
(UYWA-) UNIV WAKE FOREST.
XX
Wheeler KT, Mach RH, Childers S, Shelness G, Wang L;
WPI; 2001-662943/76.
N-PSDB; AAS15569.
XX
Novel isolated polynucleotide encoding sigma1beta receptor useful in
screening assay to identify ligands specific for the sigma1beta receptor
for tumor imaging, diagnostic and treatment methods.
XX
Claim 9; Fig 2; 56pp; English.

The invention relates to an isolated polynucleotide (I) encoding sigma 1
receptor splice variant isolated from mouse adenocarcinoma cells and
human breast tumour cells (I) or the encoded protein (II) is useful for
screening compounds useful in the imaging and treatment of proliferative
masses (i.e. tumours) and in the non-invasive diagnosis of cancer,
preferably in the diagnosis of proliferative cancer cells. (I), a cell
comprising (I), or (II) is useful in screening assays to identify ligands
specific for the sigma 1 receptor, and identification of the sigma 1
receptor permits the design of tumour imaging, diagnostic and treatment
methods. (I) is useful for preparing (II), which is useful as immunogen
for making antibodies. These antibodies are useful for a variety of
diagnostic and imaging purposes. Sigma 1 ligands obtained from screening

CC are useful as diagnostic compounds for imaging of, for example, tumour
CC cells, for determining the proliferative status of the tumour, and as
CC therapeutics for the treatment of cancer and other disorders of cell
CC proliferation. The present sequence represents the amino acid sequence of
CC human sigma 1 receptor splice variant, sigma1beta, as described in the
CC invention
XX
SQ

Sequence 192 AA;

Query Match 100.0%; Score 1016; DB 4; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQWAVGRWMAALLLAVAAVLTVVWLWLGTSQSFVQREIEIAQLARQYAGLDHLEAFSR 60
DB 1 MQWAVGRWMAALLLAVAAVLTVVWLWLGTSQSFVQREIEIAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALSGRHSGET 120
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALSGRHSGET 120
QY 121 VVHGPGEATAVWGNTWVVEYGRGVPSTLAFALADTVFSTQDFLTFLTYTLRSYARGLR 180
DB 121 VVHGPGEATAVWGNTWVVEYGRGVPSTLAFALADTVFSTQDFLTFLTYTLRSYARGLR 180
QY 181 LELTYLFGQDP 192
DB 181 LELTYLFGQDP 192

RESULT 2

AAV03755
ID AAY03755 standard; protein; 223 AA.

AC AAY03755;
XX

DT 10-JUN-1999 (first entry)

XX Human sigma receptor (SIGR) polypeptide.

XX Sigma receptor; SIGR; vesicle-trafficking disorder; gastrointestinal;
KW inflammatory; immunological disorder; AIDS; arthritis; infection; cancer;
KW neoplastic disorder.

XX Homo sapiens.

XX US5863766-A.

XX 26-JAN-1999.

XX 12-SEP-1997; 97US-00928612.

XX 12-SEP-1997; 97US-00928612.

XX (INCY-) INCYTE PHARM INC.

XX Corley NC, Hillman JL, Shah P;

XX WPI; 1999-253089/21.

XX N-PSDB; AAX29478.

XX New DNA encoding human sigma receptor polypeptide - useful for treating
PT or preventing vesicle-trafficking, immunological and neoplastic
PT disorders.

XX Claim 1; Fig 1A-D; 28pp; English.

XX This represents a human sigma receptor (SIGR) polypeptide. Host cells
CC containing a vector comprising the SIGR nucleic acid are used for the
CC recombinant expression of the protein. Recombinant SIGR can be used to
CC treat or prevent vesicle-trafficking disorders, e.g. cystic fibrosis,
CC glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes
CC insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's

CC disease, and Addison's disease; gastrointestinal disorders (ulcerative
CC colitis, gastric and duodenal ulcers); and other conditions associated
CC with abnormal vesicle trafficking (e.g. hay fever, urticaria (hives),
CC inflammatory bowel disease, myasthenia gravis, Chediak-Higashi syndrome,
CC systemic lupus erythematosus, toxic shock syndrome, and traumatic tissue
CC damage). SIGR can be used to treat or prevent immunological disorders
CC such as AIDS, adult respiratory distress syndrome, allergies, anemia,
CC asthma, atherosclerosis, bronchitis, cholecystitis, atopic dermatitis,
CC Crohn's disease, dermatomyositis, diabetes mellitus, glomerulonephritis,
CC gout, hyperesoinophilia, lupus erythematosus, multiple sclerosis,
CC myasthenia gravis, myocardial or pericardial inflammation, rheumatoid
CC arthritis, osteoarthritis, osteoporosis, pancreatitis, polymyositis,
CC scleroderma, Sjogren's syndrome, Werner syndrome, autoimmune thyroiditis;
CC complications of cancer, hemodialysis, and extracorporeal circulation; and
CC parasitic, viral, bacterial, fungal, helminth, and protozoal infections.
CC SIGR may also be used to treat or prevent neoplastic disorders e.g.
CC adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma,
CC teratocarcinoma, and cancers of the adrenal gland, bladder, ganglia,
CC gastrointestinal tract, heart, breast, cervix, liver, lung, muscle,
CC ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin,
CC spleen, testis, thymus, thyroid, and uterus
XX

SQ Sequence 223 AA;

Query Match 97.5%; Score 990.5; DB 2; Length 223;

Best Local Similarity 86.1%; Pred. No. 1.1e-101;

Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;

QY 1 MQWAVGRWMAALLLAVAAVLTVVWLWLGTSQSFVQREIEIAQLARQYAGLDHLEAFSR 60

DB 1 MQWAVGRWMAALLLAVAAVLTVVWLWLGTSQSFVQREIEIAQLARQYAGLDHLEAFSR 60

QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALSGRHS --- 117

DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVLLFGTALSGRHSGRY 120

QY 118 -----GETVHGPGEATAVWGNTWVVEYGRGVPSTLAFALADTVFSTQDFLTFLTYTLRSYARGLR 149

DB 121 WAEISDTTISGTFHOWREGTTKSEVPYGETVHGPGEATAVWGNTWVVEYGRGVPSTLAFALADTVFSTQDFLTFLTYTLRSYARGLR 180

QY 150 TLAFALADTVFSTQDFLTFLTYTLRSYARGLRLELTLYLFGQDP 192

DB 181 TLAFALADTVFSTQDFLTFLTYTLRSYARGLRLELTLYLFGQDP 223

RESULT 3

AAB19077

ID AAB19077 standard; protein; 223 AA.

XX AAB19077;

AC AAB19077;

XX 08-FEB-2001 (first entry)

XX Amino acid sequence of a human sigma receptor binding protein.

XX Human; sigma receptor binding protein; SRBP; cellular marker;

XX breast cancer; HSI; human steroid isomerase.

XX Homo sapiens.

XX WO200057188-A2.

XX 28-SEP-2000.

XX 14-MAR-2000; 2000WO-FR000607.

XX 19-MAR-1999; 99FR-00003418.

XX (SNFI) SANOFI-SYNTHELABO.

XX Casellas P, Simony-Lafontaine J;

XX WPI; 2000-594612/56.

XX Diagnostic composition for detecting breast cancer, useful for predicting
PT response to treatment, contains signal receptor binding protein and/or
PT human sterol isomerase.

XX Example; Fig 4; 27pp; French.

XX The present sequence represents a human signal receptor binding protein
CC (SRBP). The protein is a cellular marker, which may be used in the
CC compositions of the invention. The specification describes a diagnostic
CC composition for detecting breast cancer. The compositions comprises a
CC cellular marker, a protein that has at least 40% homology (at the amino
CC acid level) with SRBP and/or a protein that has at least 40% analogy with
CC HSI (human sterol isomerase). The compositions, or a similar composition
CC containing antibodies specific SRBP and/or HSI, is used to detect breast
CC cancer, whether hormone dependent or not. Analysis of the distribution of
CC the two markers can be used to predict responses to treatment, including
CC identification of patients at high risk of relapse or those who could
CC benefit from auxiliary treatments

XX Sequence 223 AA;

Query Match 97.5%; Score 990.5; DB 3; Length 223;
Best Local Similarity 86.1%; Pred. No. 1.1e-101;
Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;
QY 1 MQWAGRRWAAALLLAAVAVLTQVWMLGTQSFVQREIEIAQLARQYAGLDHLEAFSR 60
DB 1 MQWAGRRWAAALLLAAVAVLTQVWMLGTQSFVQREIEIAQLARQYAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGSRGHS --- 117
DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGSRGHS 120
QY 118 -----GETVVHGPGEATAVWGPNMTMWVEYGRGVIPS 149
DB 121 WAEISDTIISGTFHQWREGTKSEVFPGETVHVGPGGEATAVWGPNMTMWVEYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFLTFLFYTLRSYARGLRLELTLYLFGQDP 192
DB 181 TLAFALADTVFSTQDFLTFLFYTLRSYARGLRLELTLYLFGQDP 223

RESULT 4
ADE54607
ID ADE54607 standard; protein; 223 AA.

XX ADE54607;

XX 29-JAN-2004 (first entry)

XX Human Protein NP_005857, SEQ ID NO 412.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX W02003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GCHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

XX GENBANK; NP_005857.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 223 AA;

Query Match 97.5%; Score 990.5; DB 7; Length 223;
Best Local Similarity 86.1%; Pred. No. 1.1e-101;
Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;

QY 1 MQWAGRRWAAALLLAAVAVLTQVWMLGTQSFVQREIEIAQLARQYAGLDHLEAFSR 60

DB 1 MQWAGRRWAAALLLAAVAVLTQVWMLGTQSFVQREIEIAQLARQYAGLDHLEAFSR 60

QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGSRGHS --- 117

DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGSRGHS 120

QY 118 -----GETVVHGPGEATAVWGPNMTMWVEYGRGVIPS 149

DB 121 WAEISDTIISGTFHQWREGTKSEVFPGETVHVGPGGEATAVWGPNMTMWVEYGRGVIPS 180

QY 150 TLAFALADTVFSTQDFLTFLFYTLRSYARGLRLELTLYLFGQDP 192

DB 181 TLAFALADTVFSTQDFLTFLFYTLRSYARGLRLELTLYLFGQDP 223

RESULT 5

ABM81032

ID ABM81032 standard; protein; 223 AA.

XX ABM81032;

XX 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) polypeptide PRO36808, SEQ:2671.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;

XX tumour; diagnosis; cell proliferative disorder; breast cancer;

XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;

KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.

XX Homo sapiens.

OS WO2004030615-A2.

PN 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

XX 02-OCT-2002; 2002US-0411497P.

XX (GETH) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

DR N-PSDB; ACN38882.

XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

XX Claim 12; SEQ ID NO 2671; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention

XX Sequence 223 AA;

SQ Query Match 97.5%; Score 990.5; DB 8; Length 223;
Best Local Similarity 86.1%; Pred. No. 1.1e-101;
Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;

QY 1 MQWAVGRWAWAALLLAVAAVLTVQVWLWLTQSFVQREETAQLARQVAGLDHELAFSR 60

DB 1 MQWAVGRWAWAALLLAVAAVLTVQVWLWLTQSFVQREETAQLARQVAGLDHELAFSR 60

QY 61 LIVELRRLHPGHVLPDEELQWVFNAGGWMGACMLLHASLSYVLLFGTALSGRHS--- 117

DB 61 LIVELRRLHPGHVLPDEELQWVFNAGGWMGACMLLHASLSYVLLFGTALSGRHSGRY 120

QY 118 -----GETVHVGPGEATAVEMGPNMTWVYGRGVIPS 149

DB 121 WAEISDTIISCTFHQWREGTKSEVPYPGETVHVGGEATAVEMGPNMTWVYGRGVIPS 180

QY 150 TLAFALADTVFSTQDFLTLYTLRSYARGRLLELTLYLFGQDP 192

DB 181 TLAFALADTVFSTQDFLTLYTLRSYARGRLLELTLYLFGQDP 223

RESULT 6

ADE54605

ID ADE54605 standard; protein; 223 AA.

XX ADE54605;

XX 29-JAN-2004 (first entry)

XX Rat Protein AAD01198, SEQ ID NO 410.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AAD01198.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 223 AA;

Query Match 92.0%; Score 934.5; DB 7; Length 223;

Best Local Similarity 79.8%; Pred. No. 1.9e-95;

Matches 178; Conservative 4; Mismatches 10; Indels 31; Gaps 1;

QY 1 MQWAVGRWAWAALLLAVAAVLTVQVWLWLTQSFVQREETAQLARQVAGLDHELAFSR 60

Db 1 MPWAGRRWAWITLFTITVAVLQAVWMLGTQSFVQREETIAQLARQVAGLDHLEAFSR 60
61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHSLSEYVLLFGTALGSRGHS--- 117
61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHSLSEYVLLFGTALGSRGHSRY 120
118 -----GETVVHGPGEATAVWGPNNTWVVEYGRGVIPS 149
121 WAEISDTISGTHQWREGTKSEVYYPGETVVHGPGEATDVWGPNNTWVVEYGRGVIPS 180
150 TLAPALADTVFSTQDFLTFLPYTLRSYARGRLRLLELTLYLFGQDP 192
181 TLAFALSDTIFSTQDFLTFLPYTLRAVARGRLRLLELTLYLFGQDP 223

RESULT 7
AAU09155
ID AAU09155 standard; protein; 192 AA.
AC AAU09155;
XX
XX
XX 29-JAN-2002 (first entry)
XX Mouse sigma 1 receptor splice variant, sigma1beta.
XX
XX Mouse; sigma 1 receptor; sigma1beta; tumour imaging; cancer;
KW cell proliferation disorder; tumour; diagnostic; cytostatic.
XX
XX Mus sp.
XX
XX W0200174297-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010650.
XX
XX 31-MAR-2000; 2000US-0193694P.
XX
XX (UYWA-) UNIV WAKE FOREST.
XX
XX Wheeler KT, Mach RH, Childers S, Shelness G, Wang L;
XX
XX WPI; 2001-662943/76.
XX
XX N-PSDB; AAS15570.

Novel isolated polynucleotide encoding sigma1beta receptor useful in screening assay to identify ligands specific for the sigma1beta receptor for tumor imaging, diagnostic and treatment methods.

Claim 21; Fig 4; 56pp; English.

The invention relates to an isolated polynucleotide (I) encoding sigma 1 receptor splice variant isolated from mouse adenocarcinoma cells and human breast tumour cells. (I) or the encoded protein (II) is useful for screening compounds useful in the imaging and treatment of proliferative masses (i.e. tumours) and in the non-invasive diagnosis of cancer, CC preferably in the diagnosis of proliferative cancer cells. (II), a cell CC comprising (I), or (II) is useful in screening assays to identify ligands CC specific for the sigma 1 receptor, and identification of the sigma 1 CC receptor permits the design of tumour imaging, diagnostic and treatment CC methods. (I) is useful for preparing (II), which is useful as immunogen CC for making antibodies. These antibodies are useful for a variety of CC diagnostic and imaging purposes. Sigma 1 ligands obtained from screening CC are useful as diagnostic compounds for imaging of, for example, tumour CC cells, for determining the proliferative status of the tumour, and as CC therapeutics for the treatment of cancer and other disorders of cell CC proliferation. The present sequence represents the amino acid sequence of CC mouse sigma 1 receptor splice variant, sigma1beta, as described in the CC invention

Sequence 192 AA;

SQ

Query Match 91.0%; Score 925; DB 4; Length 192;
Best Local Similarity 89.0%; Pred. NO. 1.8e-94;
Matches 170; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 1 MOWAVGRRWAAALLLAVAAVLTVVWMLGTQSFVQREETIAQLARQVAGLDHLEAFSR 60
DB 1 MPWAGRRWAWITLFTITVAVLQAAWMLGTQNFVFSREETIAQLARQVAGLDHLEAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHSLSEYVLLFGTALGSRGHSGET 120
DB 61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHSLSEYVLLFGTALGSRGHSGET 120
QY 121 VVHGPGEATAVWGPNNTWVVEYGRGVIPSTLAPALADTVFSTQDFLTFLPYTLRSYARGRLR 180
DB 121 VVHGPGEATALEWGPNTWVVEYGRGVIPSTLFPALADTVFSTQDFLTFLPYTLRAYARGRLR 180
QY 181 LELTLYLFGQD 191
DB 181 LELTLYLFGQD 191

RESULT 8
AAB58747
ID AAB58747 standard; protein; 195 AA.
XX
XX AAB58747;
XX
XX 27-MAR-2001 (first entry)
XX
XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 455.

Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neurotropic; neurproliferative; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

Homo sapiens.

W0200055173-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US005881.

12-MAR-1999; 99US-0124270P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-611515/58.

N-PSDB; AAF21650.

New human breast and ovarian cancer associated gene sequences and the PT polypeptides encoded by these genes, useful in the prevention, treatment PT and diagnosis of cancer, immune disorders, cardiovascular disorders and PT neurological diseases.

Claim 11; Page 883-884; 1299pp; English.

Sequences AAF21614 - AAF22031 represent DNA sequences encoding human CC proteins AAB58711 - AAB59128. The DNA and protein sequences are CC associated with breast and ovarian cancer. Included in the invention are CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the CC isolation and characterisation of the DNA and protein sequences of the CC invention. The breast and ovarian cancer associated DNA, protein, agonist CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic; CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; CC antiinflammatory; antiulcer; vulnary; anticonvulsant; antibacterial;

CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases
 XX
 XX Sequence 195 AA;

Query Match 81.4%; Score 827.5; DB 3; Length 195;
 Best Local Similarity 83.9%; Pred. No. 1.4e-83;
 Matches 162; Conservative 0; Mismatches 0; Indels 31; Gaps 1;
 QY 31 GTQSFVQREETAQLARQAGLDHELAFSLRLVELRLHPGHVLPDEELQWVFNAGGWM 90
 DB 3 GTQSFVQREETAQLARQAGLDHELAFSLRLVELRLHPGHVLPDEELQWVFNAGGWM 62
 QY 91 GAMCLLHASLSYVLLFGTALSGRHS-----GE 119
 DB 63 GAMCLLHASLSYVLLFGTALSGRHSGRYWAIEISDTIISGTFHQWREGTKSEVFYDGE 122
 QY 120 TVVHGPGETAWEVGNPTWVYGRGVPSTLAFALADTVFSTQDFTLFTYLRSYARGL 179
 DB 123 TVVHGPGETAWEVGNPTWVYGRGVPSTLAFALADTVFSTQDFTLFTYLRSYARGL 182
 QY 180 RLELTTYLFGQDP 192
 DB 183 RLELTTYLFGQDP 195

RESULT 9
 ABB06121
 ID ABB06121 standard; protein; 106 AA.
 AC ABB06121;
 XX
 XX 10-MAY-2002 (first entry)
 XX
 XX Human NS protein sequence SEQ ID NO:213.
 XX
 KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
 KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
 KW anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuroleptic;
 KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200206315-A2.
 XX
 XX 24-JAN-2002.
 XX
 XX 17-JUL-2001; 2001WO-IL000653.
 XX
 XX 18-JUL-2000; 2000IL-00137345.
 XX
 XX 15-DEC-2000; 2000IL-00140354.
 XX
 XX (COMP-) COMPUGEN LTD.
 XX

PI Mintz L, Freilich S, Bernstein J;
 XX WPI; 2002-155037/20.
 DR N-PSDB; ABL39775.
 XX
 PT One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 XX
 XX Claim 6; Page 248; 290pp; English.
 XX
 XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC anorectic, muscular, antiinfertility, cardiovascular, dermatological,
 CC vasotropic, antiarteriosclerotic, antiinflammatory, cardiovascular,
 CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
 CC immunomodulator, anticonvulsant, antidiabetic, antiasthmatic, cardiant,
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC nootropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX
 XX Sequence 106 AA;

Query Match 53.5%; Score 544; DB 5; Length 106;
 Best Local Similarity 98.1%; Pred. No. 2.4e-52;
 Matches 103; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MCWAVGRWAWAALLLVAVALTQVWLWLTQSFVFOREETAQLARQAGLDHELAFSR 60
 DB 1 MCWAVGRWAWAALLLVAVALTQVWLWLTQSFVFOREETAQLARQAGLDHELAFSR 60
 QY 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
 DB 61 LIVELRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSYVL 105
 RESULT 10
 AAR90670
 ID AAR90670 standard; protein; 222 AA.
 XX AAR90670;
 XX
 XX 25-MAR-2003 (revised)
 DT 11-JUL-1996 (first entry)
 XX
 XX S. cerevisiae delta 8-7 isomerase.
 DE
 XX Delta 8-7 isomerase; sterol; biosynthesis; accumulation; composition;
 KW modulation; mutant; yeast; enzyme.
 XX
 XX Saccharomyces cerevisiae.
 OS
 XX US5480805-A.
 XX
 XX 02-JAN-1996.
 XX
 XX 10-MAY-1994; 94US-00240496.
 XX
 XX 12-AUG-1992; 92US-00929764.
 XX
 XX (STAD) AMOCO CORP.
 XX
 XX Wolf FR, Cuellar RE;

XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 1902; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 222 AA;
Query Match 20.9%; Score 212.5; DB 8; Length 222;
Best Local Similarity 30.0%; Pred. No. 5.3e-15; Indels 43; Gaps 8;
Matches 62; Conservative 36; Mismatches 66;
QY 14 LLLAAVAALTVQVW-LWLGTQSFVFOREIAQLAQYAGLDHE---LAFSLRILVLRLLH 69
DB 8 LLIGVGVYIMVFLTWLFT-NYMPDPKTLNEICNSVISKHNAEGLSTEDLLQVDRDL 66
QY 70 PGHVLDPDEL-----QWVFVAGGMWGMACLLHLSLSEYVLLFGTALGSRHSGE---- 119
DB 67 ASH-YGDEYINRYVKEEWFVFNAGGAGQMIIILHASVSEYLLIFGTAVGTEGHTGVHFD 125
QY 120 ---TVVHG-----PGEATAVEWG-----PNTWMVEYGRGVIPSTILA 152
DB 126 DYFTILHGTQIAALPYATEAEVYTPGTHHLKKYAKQYMSPGGSPALSLAQGWIPCMPLP 185
QY 153 FALADTVFSTQDFTLFLYTLRSYARGL 179
DB 186 FGFLDTFSSTLDLYTLRYVTLTARDM 212
RESULT 13
ADN19490
ID ADN19490 standard; protein; 219 AA.
XX
XX ADN19490;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #2143.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polypeptide.
XX
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2143; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 219 AA;
Query Match 20.3%; Score 206; DB 8; Length 219;
Best Local Similarity 32.9%; Pred. No. 2.7e-14; Indels 38; Gaps 4;
Matches 55; Conservative 22; Mismatches 52;
QY 39 REEIAQLAQYAGLDHELAFLSRLIVELRLRHGHVLPDEELQWVFVAGGMWGMACLLHA 98
DB 42 QKSIALYANDTKALLYDLS-DELVAEYGDL-----ITPVNQDEWHNNAGGAMGTWFLHA 96
QY 99 SLSEYVLLFGTALGSRHSG-----ETVHGPGGATAVEWGP 135
DB 97 SPSEYLIFFGTPIGTGEGHSGVHMADDYFTILRQLAASANDLEARVLPDGHVHPWGH 156
QY 136 NT-----WMVEYGRGVIPSTILAFLADTVFSTQDFTLFLYTL 172
DB 157 TQAYSMPSGPECFALSLAQGWIMVSLPFGFMDGLFSTIDFGTLYKTV 203

RESULT 14

Query Match	20.1%	Score 204;	DB 8;	Length 285;
Best Local Similarity	29.5%;	Pred. No. 6.5e-14;		
Matches	69:	Conservative	33:	Mismatches 62:
				Indels 70:
				Gaps 8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 04:41:20 ; Search time 42 Seconds
(without alignments)
341.253 Million cell updates/sec

Title: US-09-823-069A-2
Perfect score: 1016
Sequence: 1 MQMVGRRWAAALLAVAA.....RSVARGRLLELTYYLFGQDP 192

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/6CTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	990.5	97.5	223	2	US-08-928-612-1
2	918.5	90.4	223	2	US-08-928-612-3
3	212.5	20.9	222	1	US-08-240-496A-2
4	83.5	8.2	238	4	US-09-502-540-15642
5	83	8.2	194	1	US-08-063-552-8
6	83	8.2	194	5	PCT-US93-05704-8
7	82	8.1	1045	4	US-09-328-352-5426
8	81.5	8.0	441	4	US-09-328-352-5426
9	81.5	8.0	487	4	US-09-328-352-6767
10	81	8.0	463	4	US-09-328-352-6767
11	80.5	7.9	353	3	US-09-461-474-4
12	80.5	7.9	494	4	US-09-252-991A-19867
13	77	7.6	317	4	US-09-489-039A-10626
14	76.5	7.5	705	4	US-09-328-352-4457
15	76	7.5	376	4	US-09-252-991A-19381
16	75.5	7.4	509	4	US-09-270-767-41638
17	74.5	7.3	623	4	US-09-252-991A-19867
18	74	7.3	162	4	US-09-902-540-11988
19	74	7.3	480	4	US-09-252-991A-18961
20	73.5	7.2	252	4	US-09-328-352-7284
21	73.5	7.2	295	4	US-09-252-991A-20718
22	73.5	7.2	551	4	US-09-489-039A-14227
23	73.5	7.2	1165	4	US-09-949-016-6874
24	73	7.2	295	4	US-09-583-110-3181
25	73	7.2	323	4	US-09-107-433-3410
26	73	7.2	335	4	US-09-602-787A-518
27	73	7.2	398	4	US-09-252-991A-19152

28	73	7.2	527	4	US-09-602-787A-516
29	72.5	7.1	391	4	US-09-543-681A-7029
30	72.5	7.1	1165	4	US-09-949-016-11392
31	72.5	7.1	7257	3	US-09-335-409-5
32	72.5	7.1	7257	3	US-09-568-102-5
33	72.5	7.1	7257	3	US-09-567-963-5
34	72.5	7.1	7257	3	US-09-568-480-5
35	72.5	7.1	7257	3	US-09-568-486-5
36	72.5	7.1	7257	3	US-09-568-472-5
37	72.5	7.1	7257	3	US-09-567-899-5
38	71.5	7.0	178	4	US-09-543-681A-7183
39	71.5	7.0	460	4	US-09-252-991A-24308
40	71.5	7.0	1996	2	US-08-804-227C-9
41	71.5	7.0	1996	2	US-08-804-198-3
42	71	7.0	373	4	US-09-622-439-6
43	71	7.0	373	4	US-09-622-439-26
44	71	7.0	373	4	US-10-318-142-6
45	71	7.0	373	4	US-10-318-142-26

ALIGNMENTS

RESULT 1
US-08-928-612-1
; Sequence 1, Application US/08928612
; Patent No. 5863766
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGMA RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,612
; FILING DATE: HEREMITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION/DOCKET NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0383 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT04
; CLONE: 1542751
; US-08-928-612-1
Query Match 97.5%; Score 990.5; DB 2; Length 223;
Best Local Similarity 86.1%; Pred. No. 4e-106;

Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;
QY 1 MQWAVGRWAWAALLAVAVLTQVWVLWGTSQVFFOREEIAQLARQVAGLDHLEAFSR 60
Db 1 MQWAVGRWAWAALLAVAVLTQVWVLWGTSQVFFOREEIAQLARQVAGLDHLEAFSR 60
QY 61 LIVELRRLHPGHVLPDEELQWVFNAGGMMGAMCILLHASLSLEYVLLFGTALGSRGHS--- 117
Db 61 LIVELRRLHPGHVLPDEELQWVFNAGGMMGAMCILLHASLSLEYVLLFGTALGSRGHSRY 120
QY 118 -----GETVVHGPGEATAVEMGPNMTMMVEYGRGVIPS 149
Db 121 WAEISDTIISGTFHQWREGTTKSEVFPYGETVVHGPGEATAVEMGPNMTMMVEYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFLFTLYTLRSYARGRLRLLELTLYLFGQDP 192
Db 181 TLAFALADTVFSTQDFLFTLYTLRSYARGRLRLLELTLYLFGQDP 223

RESULT 2
US-08-928-612-3
; Sequence 3, Application US/08928612
; Patent No. 5863766
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SIGMA RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,612
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0383 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1403300
US-08-928-612-3

Query Match 90.4%; Score 918.5; DB 2; Length 223;
Best Local Similarity 79.4%; Pred. No. 8.2e-98;
Matches 177; Conservative 4; Mismatches 11; Indels 31; Gaps 1;
QY 1 MQWAVGRWAWAALLAVAVLTQVWVLWGTSQVFFOREEIAQLARQVAGLDHLEAFSR 60

Db 1 MQWAVGRWAWAALLAVAVLTQVWVLWGTSQVFFOREEIAQLARQVAGLDHLEAFSK 60
QY 61 LIVELRRLHPGHVLPDEELQWVFNAGGMMGAMCILLHASLSLEYVLLFGTALGSRGHS--- 117
Db 61 LIVELRRLHPGHVLPDEELQWVFNAGGMMGAMCILLHASLSLEYVLLFGTALGSRGHSRY 120
QY 118 -----GETVVHGPGEATAVEMGPNMTMMVEYGRGVIPS 149
Db 121 WAEISDTIISGTFHQWREGTTKSEVFPYGETVVHGPGEATAVEMGPNMTMMVEYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFLFTLYTLRSYARGRLRLLELTLYLFGQDP 192
Db 181 TLAFALADTVFSTQDFLFTLYTLRSYARGRLRLLELTLYLFGQDP 223

RESULT 3
US-08-240-496A-2
; Sequence 2, Application US/08240496A
; Patent No. 5480805
; GENERAL INFORMATION:
; APPLICANT: Wolf, Fred R.
; APPLICANT: Cuellar, Richard E.
; TITLE OF INVENTION: Composition for Modulating Sterols in
; TITLE OF INVENTION: Yeast
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Joanne M. Giesser
; STREET: Amoco Corp., Suite 600, 55 Shuman Blvd.
; CITY: Naperville
; STATE: IL
; COUNTRY: USA
; ZIP: 60563
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,496A
; FILING DATE: 10-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/929,764
; FILING DATE: 12-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Giesser, Joanne M.
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 31,456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 717-2443
; TELEFAX: (708) 717-2430
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-240-496A-2

Query Match 20.9%; Score 212.5; DB 1; Length 222;
Best Local Similarity 30.0%; Pred. No. 2.7e-16;
Matches 62; Conservative 36; Mismatches 66; Indels 43; Gaps 8;
QY 14 LILAVAAVLTQVWV-LWLGTQSFVQREIEIAQLARQVAGLDHE---LAFSRLIVELRLH 69
Db 8 LLIGVGVYIMNVLTFTTLPT-NYFDPKTLNEICNSVLSKNAASGLSTEDLLQDVRDAL 66
QY 70 PGHVLDPDEL-----QWVFNAGGMMGAMCILLHASLSLEYVLLFGTALGSRGHS--- 119
Db 67 ASH-YGDEYINRYKKEWVFNAGGMMQMIILHASVSEYILFCTAVGTEGTVGFAD 125

QY 120 -----TVVHG-----PGEATAVEMG-----PNTVMVEYGRGVIPSTIA 152
Db 126 DVFTHLGTQIAALPYATAEAVYTPGTHLKKGYAKQYMPGSGFALEAGWIPCMPL 185
QY 153 FALADTVFSTODFLTYLTYLSYARGL 179
Db 186 FGLDTFSSTLDTLYTLTYRTVYLTARDM 212

RESULT 4
US-09-902-540-15642
; Sequence 15642, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15642
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15642

Query Match 8.2%; Score 83.5; DB 4; Length 298;
Best Local Similarity 27.7%; Pred. No. 0.33; 88; Indels 49; Gaps 11;
Matches 61; Conservative 22; Mismatches 22

QY 4 AVGRRWAAALLAVAAVLTVVWLMLGTQSFVFOR--BEIAQARQYAGLDH--ELAFSLI--VELRRLHP 70
Db 33 ANDRAVSLAALQA--AAPESKAPAHHLGT---LLSRAVDPVLSRLHA--RDVALALSCA 87
QY 60 -----RLIVELRRLHPGHVLPDEELQWVFN-----AGQWMA 92
Db 88 QGLPAARALLQEQVLLKRLVPLRIHPSPVFADEVLQALRANLLMPRAEAPSLRLGYAGA 147
QY 93 MCLLH---ASLSEYVLLFGTALGSRGH--SGETVHVGPGGATAVEMGPNMTVMVEYGRGVIP 148
Db 148 GPLLHWVSISATRLALMRKALGASHVEAVLAHPAG--GUELG---FVREEARGHVR 203
QY 149 STLAFALADTVFSTODFLTYLTYLSYARGLRLLELTYLF 188
Db 204 AAFVRAVSLDDEDRLLRLHFVER-----LSLERMGALF 238

; APPLICATION NUMBER: US/08/063,552
; FILING DATE: 19930514
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9067-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Transposon 10
; US-08-063-552-8

Query Match 8.2%; Score 83; DB 1; Length 194;
Best Local Similarity 29.2%; Pred. No. 0.2; 42; Indels 32; Gaps 7;
Matches 38; Conservative 18; Mismatches 42

QY 14 LLLAVAAVLTVVWLMLGTQSFVFOR--BEIAQARQYAGLDH--ELAFSLI--VELRRLHP 70
Db 45 VLLALYALMOVIFAPWLGKMSDRFGRPVLLSLIGASLDYLLAFSSALMMLYLGRLLS 104
QY 71 G-----HVLPD-----BELOVFNAGGWMGMCLLHASLSEYVLLFTALGSRG 115
Db 105 GITGATGAVAASVIADTTTSASQKVKWF-----GWLGSFGLG-----LIAGPIIG--G 150
QY 116 HSGETVHGP 125
Db 151 FAGEISPHSP 160

RESULT 6
PCT-US93-05704-8
; Sequence 8, Application PC/TUS9305704
; GENERAL INFORMATION:
; APPLICANT: Edwards, Robert H
; TITLE OF INVENTION: Vesicle Membrane Transport Proteins
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05704
; FILING DATE: 19930611
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9067-1PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: AMINO ACID

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Transposon 10
PCT-US93-05704-8

Query Match 8.2%; Score 83; DB 5; Length 194;
Best Local Similarity 29.2%; Pred. No. 0.2;
Matches 38; Conservative 18; Mismatches 42; Indels 32; Gaps 7;

QY 14 LLLAAVAVLTQVWMLGTSQVFOREEIAQLARQYAGLDH-ELAFSRLI--VELRLRHP 70
DB 45 VLLALYALMQVIFAPWLGKMSDRFGRFPVLLSLIGASLDYLLAFSALWMLYIGRLLS 104

QY 71 G-----HVLDP-----EELQWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGRG 115
DB 105 GITGATGAVASVIADTTTSASQKVWF-----GWLGSFGLG-----LIAGPIIG--G 150

QY 116 HSGETVVHGP 125
DB 151 PAGEISPHSP 160

RESULT 7
US-09-252-991A-17661
; Sequence 17661, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17661
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17661

Query Match 8.1%; Score 82; DB 4; Length 1045;
Best Local Similarity 32.6%; Pred. No. 2.9;
Matches 42; Conservative 13; Mismatches 50; Indels 24; Gaps 7;

QY 10 ANAALLAV-----AAVLTVVWLW---LGTQSFVFOREEIAQLARQYAGLDH-ELAFSRLI 62
DB 356 SWSALLDALARLGAARVTTQVDLLATRAGGQDHALGHPELHARQVGHIERQAAAFQGAR 415

QY 63 VELRLRHPG-HV---LPDEELQWVFNAGWGMGAMCLLHASLSEYVLL-FGTALSGRHS 117
DB 416 VGRALDPGEHVTLRVADVQRQ-----AQLVGA-----FDEFALLDQGDQAVDRGES 463

QY 118 GETVVHGP 126
DB 464 VEVDLRGDG 472

RESULT 8
US-09-328-352-5426
; Sequence 5426, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5426
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5426

Query Match 8.0%; Score 81.5; DB 4; Length 441;
Best Local Similarity 26.1%; Pred. No. 0.98;
Matches 42; Conservative 27; Mismatches 67; Indels 25; Gaps 8;

QY 40 EEIAQLARQYAGLDH-ELAF-----SRLIVELRLHPGHVLPDEELQWVFNAGW 89
DB 232 EGISKITAQDVKYAEELGFRIKHGIARRAEKGIEL-RVHP-TLIPDQL-----IANVGV 286

QY 90 MGAMCLLHASLSEYVLLFGTALSGRHSGETVVHGPGEATAVEWGPNTWVVEYGRGVIPS 149
DB 287 KNAV-LVQANAVGPTLYYGAGAGA-GPTASAVV-----ADVIDIVRDISYTDGAGTIPQ 339

QY 150 TLAFALADTVFSTQDFLFTYTLRSYAR---GLRLLETTYL 187
DB 340 LAPEALTNPILSRREMTGGYIRLNAEDDOTGVLADVTTL 380

RESULT 9
US-09-328-352-6767
; Sequence 6767, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6767
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6767

Query Match 8.0%; Score 81.5; DB 4; Length 487;
Best Local Similarity 21.0%; Pred. No. 1.1;
Matches 37; Conservative 30; Mismatches 66; Indels 43; Gaps 7;

QY 13 ALLLAVAAVLTVVWLWLTQ-----SFVFOREEIAQLARQYAGLDH-ELA 57
DB 33 ATSIGLACFLAVLEWKWLRTQNPYKDLFKYWKIPAFVAFGVGVVSVYQFGTNWSE 92

QY 58 FSRILVELRLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSEYVLLFGTA-LGSRGH 116
DB 93 FSRVAGSI-----TGPLLTYEVLSAFFLEA-GFLG-----IMLFGMGRVGPRAH 135

QY 117 SGETVVHGPGEATAVEW---GNTWVVEYGRGVISTLAFALADTVFSTQDFLTLEY 170
DB 136 FFATLMVAIGTCISMFWILSSNM-----QTPOGFAIENGIIVPKDWLAIVF 183

RESULT 10
US-09-252-991A-19381
; Sequence 19381, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19381
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19381

Query Match      8.0%; Score 81; DB 4; Length 463;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 53; Conservative 27; Mismatches 85; Indels 62; Gaps 9;

QY 2 QNVA---GRWAAALLAAVLTQVWMLGTQSFVQREBIAOLA---ROYAGLDHE 55
DB 230 QMLADGGHPAVPATLAVLLASLALVVPAMPETRLAGTPPATLAFRRVLA--DRP 287

QY 56 LAFSRLIVELRLRHGHPVLPDELOWVFNAG-----GWMGAMCLLHASLSEY 103
DB 288 LQTRALLVAVLN-----LVFSFYAAGPFPMVGDLPGLGFGWIGLAIAIAGSLGAL 337

QY 104 V-----LLFGTALSGRHSGETVHVGPGEATAVWGPNTVMVEYGRGV-IPS 149
DB 338 LNRRLPRTWNSARRVRLGLALAAAGATAQTLLAAVGAEGLYWALPALPIFIFGVAIDN 397

QY 150 TLAFAL-----ADTFESTQDFTLFTYLSYARGRLLELTYYL 187
DB 398 LGLPALHAYDDCRGAGALFGLAYLLI-----GLGLGASTLL 435

RESULT 11
US-09-461-474-4
; Sequence 4, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafaleki, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BB1303 US NA
; CURRENT APPLICATION NUMBER: US/09/461,474
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: 60/112,562
; EARLIER FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Glycine max
US-09-461-474-4

Query Match      7.9%; Score 80.5; DB 3; Length 353;
Best Local Similarity 24.5%; Pred. No. 0.93;
Matches 34; Conservative 18; Mismatches 58; Indels 29; Gaps 4;

QY 72 HVLDPDELOWVFNAGWNG-----AMCLLHASLSEYVLLFCT-----ALGSRGHS 117
DB 11 NVLEQETLKWVFGVGKGVGKTTCCSILSILLATVRSSVLIISTDPAHNLSDAFQORFTK 70

QY 118 GETVHVGPGEATAVWGPNTVMVEYG-----RCVIP---STLAFALADTVFST 162
DB 71 TPTLVNGFSNLYAMEVDFTVEHDMGADGMDTLFSELAGAIPIGIDEAMSFAEMLKLIVQT 130

QY 163 QDFTLFTYLSYARGRL 181
DB 131 MDYSIVDPAPTGHILRL 149

RESULT 12
US-09-252-991A-32227
; Sequence 32227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32227
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (139)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-32227

Query Match      7.9%; Score 80.5; DB 4; Length 494;
Best Local Similarity 23.0%; Pred. No. 1.5;
Matches 32; Conservative 16; Mismatches 50; Indels 41; Gaps 3;

QY 14 LLLAAVAAVLTQVWMLWLTQSFVQREBIAQLARQYAGLDHDLAFSLIVELRLRHGPHV 73
DB 99 LVVVVLLALSHVV-----DEALELRMVADLERPHGQV 132

QY 74 LPDEELQWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGRHSGETVHVGPGEATAVEW 133
DB 133 -----EHVGXVXGVVGMARAGHAEVLLAVVEQAQVHAGTVELEGRAEHQGALL 182

QY 134 GPNTWMV-----EYGRGVI 147
DB 183 GRSQVVVAAGVEAQGEVV 201

RESULT 13
US-09-489-039A-10626
; Sequence 10626, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10626
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10626

Query Match      7.6%; Score 77; DB 4; Length 317;
Best Local Similarity 27.7%; Pred. No. 2;
Matches 51; Conservative 20; Mismatches 61; Indels 52; Gaps 11;

QY 22 LQGVWMLWLTQSFVQREBIAQLARQYAGLDH-----ELAFSLIVELRLRHGPHVLPD 76
DB 95 MTQYFSMW-GNLTII---RENLLFIARLYS-LDRRRRVERVALSELGLTARQ----HOLAK 145

QY 77 BELQWVFNAGWGMGAMCLLHASLSEYVLLF---GTA-----LGRGRHS 117
DB 146 E-----LSGGWKQRMALAAACMLHPEVLLFLDEPTAGVDPKARRBFWQMLHQLSDRGIS 198

QY 118 GETVHVGPGEATAVWGPNTVMVEYGRVIPSTLAFALADTVFSTQDFTLFTYLSYAR 177
```

Db 199 LLVSTHYMDEA---ERCHKVAVLSYGRLLANGTIA-----SIASQNLI-----TWRTSGA 246
QY 178 GLRL 181
Db 247 GLTL 250

RESULT 14
US-09-328-352-4457
; Sequence 4457, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4457
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4457

Query Match 7.5%; Score 76.5; DB 4; Length 705;
Best Local Similarity 26.4%; Pred. No. 7.2;
Matches 48; Conservative 23; Mismatches 56; Indels 55; Gaps 12;
QY 9 WAWAALLLAVAAVL-----TQVVLWLGTQSFVQREETAQLARQYAGLDHELAFSR 60
Db 23 WSLVALIGAISFGMLALSERGEHVNVLVLAAC-----VISTAYRF-----YSLFIAT 71
QY 61 LIVEL--RRLHPGHVLPDEELQWVFNAGGMMGAMCLLHASLSEYVLLFGTALGSRGHS 118
Db 72 KVFELNPRELTPAHLAD-GLDYVPTN-----KVV-LFGHHFAAIAGAG 113
QY 119 ETVVHGPGATAVEGPNP--WVVEYGRGVIPSTLAFALADTVFSTQDFLTFLPYTLRSYAR 177
Db 114 PLV--GPILAAQMGFLPGTILLLV--GVV-----LAGAV---QDFLVLPFISTRDGR 158
QY 178 GL 179
Db 159 SL 160

RESULT 15
US-09-252-991A-23396
; Sequence 23396, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23396
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23396

Query Match 7.5%; Score 76; DB 4; Length 376;
Best Local Similarity 26.8%; Pred. No. 3.4;
Matches 37; Conservative 12; Mismatches 49; Indels 40; Gaps 7;

QY 4 AVGRRWAAALLLAVAAVLTQV-----MLW---LGTQSFVFOREE----- 41
Db 237 APGRPWYLLLVAAVPAALLLLAVGVGLLRPELHAPAWLWQFILANLFFVSLAEEALFRGY 296
QY 42 IAQLARQYAGLDHELAFSRLIVELRRLHPGHVLPDEELQWVFNAGG-----WMGAM 93
Db 297 LQQLGQWLGPWPAALASALFGLAHFAGGPLL-----MLFAGLAGLIYGLAWLWSGRL 350
QY 94 ---CLLHASLS-EYVLLF 107
Db 351 WATLFHFGLNLTHLLLF 368

Search completed: October 29, 2005, 04:51:36
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 04:38:34 ; Search time 168 Seconds
(without alignments)
477.763 Million cell updates/sec

Title: US-09-823-069A-2
Perfect score: 1016
Sequence: 1 MQWAGRRWAAALLAVAA.....RSVAGRLRLTYLFGQDP 192

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1016	100.0	192	9	US-09-823-069-2
2	925	91.0	192	9	US-09-823-069-4
3	827.5	81.4	195	9	US-09-925-298-455
4	827.5	81.4	195	14	US-10-102-806-455
5	212.5	20.9	222	15	US-10-369-493-1902
6	206	20.3	219	15	US-10-369-493-2143
7	204	20.1	285	15	US-10-369-493-4103
8	180	17.7	259	15	US-10-369-493-13403
9	86	8.5	421	16	US-10-437-963-118957
10	85	8.5	471	16	US-10-481-265-93
11	85.5	8.4	335	14	US-10-156-761-12590

12	85	8.4	307	14	US-10-156-761-14148	Sequence 14148, A
13	83	8.2	1158	9	US-09-834-792-2	Sequence 2, Appli
14	83	8.2	1158	13	US-10-026-188-5	Sequence 5, Appli
15	83	8.2	1158	16	US-10-794-897-4	Sequence 4, Appli
16	81.5	8.0	440	15	US-10-282-122A-61930	Sequence 61930, A
17	81.5	8.0	523	15	US-10-369-493-12888	Sequence 12888, A
18	81	8.0	352	16	US-10-437-963-137218	Sequence 137218,
19	81	8.0	630	15	US-10-405-877-43	Sequence 43, Appli
20	81	8.0	840	14	US-10-156-761-13027	Sequence 13027, A
21	80.5	7.9	354	15	US-10-424-599-279814	Sequence 279814,
22	80.5	7.9	361	15	US-10-425-114-46329	Sequence 46329, A
23	80	7.9	527	17	US-10-732-923-4147	Sequence 4147, Ap
24	79.5	7.8	363	15	US-10-425-114-72368	Sequence 72368, A
25	79.5	7.8	491	15	US-10-282-122A-78405	Sequence 78405, A
26	79.5	7.8	512	16	US-10-437-963-184765	Sequence 184765,
27	79.5	7.8	1926	16	US-10-437-963-132922	Sequence 132922,
28	79	7.8	330	14	US-10-221-097-43	Sequence 43, Appli
29	78.5	7.7	469	15	US-10-369-493-16951	Sequence 16951, A
30	78	7.7	436	16	US-10-437-963-196797	Sequence 196797,
31	78	7.7	1035	15	US-10-282-122A-51321	Sequence 51321, A
32	77.5	7.6	231	16	US-10-425-115-228673	Sequence 228673,
33	77	7.6	273	17	US-10-732-923-15741	Sequence 15741, A
34	77	7.6	400	16	US-10-437-963-169852	Sequence 169852,
35	77	7.6	439	9	US-09-967-477B-4	Sequence 4, Appli
36	77	7.6	595	16	US-10-437-963-129532	Sequence 129532,
37	77	7.6	944	14	US-10-213-990-27	Sequence 27, Appli
38	77	7.6	1168	16	US-10-794-897-8	Sequence 8, Appli
39	76.5	7.5	479	14	US-10-156-761-8209	Sequence 8209, Ap
40	76	7.5	99	16	US-10-425-115-335924	Sequence 335924,
41	76	7.5	193	16	US-10-767-701-33261	Sequence 33261, A
42	76	7.5	392	18	US-10-450-763-37569	Sequence 37569, A
43	76	7.5	473	14	US-10-156-761-12103	Sequence 12103, A
44	75.5	7.4	470	17	US-10-732-923-23727	Sequence 23727, A
45	75.5	7.4	690	20	US-11-097-143-16488	Sequence 16488, A

ALIGNMENTS

RESULT 1

US-09-823-069-2
; Sequence 2, Application US/09823069
; Patent No. US20020061847A1
; GENERAL INFORMATION:
; APPLICANT: Wheeler, Kenneth
; APPLICANT: Mach, Robert
; APPLICANT: Childers, Steven
; APPLICANT: Shelnese, Gregory
; APPLICANT: Wang, Li-Ming
; TITLE OF INVENTION: METHODS AND COMPOSITIONS UTILIZING AN ALTERNATIVE SPLICE VARIANT
; FILE REFERENCE: 9151.6
; CURRENT APPLICATION NUMBER: US/09/823,069
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,694
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-069-2

Query Match 100.0%; Score 1016; DB 9; Length 192;
Best Local Similarity 100.0%; Pred. No. 7.2e-99;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQWAGRRWAAALLAVAAVLTVQVWLWLGTSFVQREIIAQLAQYAGLDHDLAFPSR 60
DB 1 MQWAGRRWAAALLAVAAVLTVQVWLWLGTSFVQREIIAQLAQYAGLDHDLAFPSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFVNAGWGMGAMCLLHASLSEYVLLFGTALGSRHSGSET 120

Db 61 LIVELRLHPGHLVLPDEELQWVFNAGGWMGAMCLLHASLSLEYVLLFGTALGSRGHSGET 120
QY 121 VVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 180
Db 121 VVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 180
QY 181 LELTTYLFGQDP 192
Db 181 LELTTYLFGQDP 192

RESULT 2

US-09-823-069-4
; Sequence 4, Application US/09823069
; Patent No. US20020061847A1
; GENERAL INFORMATION:
; APPLICANT: Wheeler, Kenneth
; APPLICANT: Mach, Robert
; APPLICANT: Childers, Steven
; APPLICANT: Shelness, Gregory
; APPLICANT: Wang, Li-Ming
; TITLE OF INVENTION: METHODS AND COMPOSITIONS UTILIZING AN ALTERNATIVE SPLICING VARIANT
; FILE REFERENCE: SIGMA-1 RECEPTOR
; CURRENT APPLICATION NUMBER: US/09/823,069
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,694
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-823-069-4

Query Match 91.0%; Score 925; DB 9; Length 192;
Best Local Similarity 89.0%; Pred. No. 3e-89;
Matches 170; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 MWAGRRWAWALLAVAVLTQVVLWLTQSFVFOREETIAQLARQYAGLDHELAFSR 60
Db 1 MWAGRRWAWTLITITIAVLIAQLWLTQSFVFOREETIAQLARQYAGLDHELAFSR 60
QY 61 LIVELRLHPGHLVLPDEELQWVFNAGGWMGAMCLLHASLSLEYVLLFGTALGSRGHSGET 120
Db 61 LIVELRLHPGHLVLPDEELQWVFNAGGWMGAMCLLHASLSLEYVLLFGTALGSRGHSGET 120
QY 121 VVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 180
Db 121 VVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 180
QY 181 LELTTYLFGQD 191
Db 181 LELTTYLFGQD 191

RESULT 3

US-09-925-298-455
; Sequence 455, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 455
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-298-455

Query Match 81.4%; Score 827.5; DB 9; Length 195;
Best Local Similarity 83.9%; Pred. No. 6e-79; Indels 31; Gaps 1;
Matches 162; Conservative 0; Mismatches 0;

QY 31 GTQSFVFOREETIAQLARQYAGLDHELAFSRLLIVELRLHPGHLVLPDEELQWVFNAGGWM 90
Db 3 GTQSFVFOREETIAQLARQYAGLDHELAFSRLLIVELRLHPGHLVLPDEELQWVFNAGGWM 62
QY 91 GAMCLLHASLSLEYVLLFGTALGSRGHS-----GE 119
Db 63 GAMCLLHASLSLEYVLLFGTALGSRGHSRYWAEISDTIISGTFHQWREGTTKSEVFPGE 122
QY 120 TVVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 179
Db 123 TVVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 182
QY 180 RLETTYLFGQDP 192
Db 183 RLETTYLFGQDP 195

RESULT 4

US-10-102-806-455
; Sequence 455, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 455
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-455

Query Match 81.4%; Score 827.5; DB 14; Length 195;
Best Local Similarity 83.9%; Pred. No. 6e-79;
Matches 162; Conservative 0; Mismatches 0; Indels 31; Gaps 1;

QY 31 GTQSFVFOREETIAQLARQYAGLDHELAFSRLLIVELRLHPGHLVLPDEELQWVFNAGGWM 90
Db 3 GTQSFVFOREETIAQLARQYAGLDHELAFSRLLIVELRLHPGHLVLPDEELQWVFNAGGWM 62
QY 91 GAMCLLHASLSLEYVLLFGTALGSRGHS-----GE 119
Db 63 GAMCLLHASLSLEYVLLFGTALGSRGHSRYWAEISDTIISGTFHQWREGTTKSEVFPGE 122
QY 120 TVVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 179
Db 123 TVVHGPGEATAVWGPNTMMVEYGRGVIPSTLAFALADTVFSTQDFLTLYTLRSYARGLR 182
QY 180 RLETTYLFGQDP 192
Db 183 RLETTYLFGQDP 195


```

; SEQ ID NO 12590
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12590

Query Match      8.4%; Score 85.5; DB 14; Length 335;
Best Local Similarity 26.0%; Pred. No. 2.8;
Matches 58; Conservative 23; Mismatches 81; Indels 61; Gaps 14;

QY 9 WAAALLLAAVLTQVWMLGT-----QSFVQREIEIAQLARQYAGLDHDLAFSRLIVE 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 WPTALLSTVAVL--LGN-WIGTHPEVQSSIAAPSELRALTR--PGGEYETVYSS---- 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 65 LRLHPGHVLPDEELQWVFVWAGWNGAMCLLHA-----SLSEYVLLFCTALGSRG 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 ----HPAASFAAQ-----VMTN-NAQAAMCLVIGIFLGLFVLWILFQNNMLNIGVGGLMS 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 116 HSQE-----TVVHGFGEATAV-----EW-----GP-----NTMVVEYGRGVIPST 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 SAGRLDTFGLVLPAGLLELTAVFVAAGTGLRGWTLIDPGPSPRTALAEGRALGWA 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 151 LAPALADTVFSTQDFTLFTLYLSYAR---GLRLELT--TYLP 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 VGLALVFSGAIEGFVTPSGLPTWARIGIGIAAEALFLTYV 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-10-156-761-14148
; Sequence 14148, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14148
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14148

Query Match      8.4%; Score 85; DB 14; Length 307;
Best Local Similarity 26.2%; Pred. No. 2.9;
Matches 51; Conservative 24; Mismatches 62; Indels 58; Gaps 9;

QY 14 LLLAAVLTQVWMLGTQSFVQREIEIAQLARQYAGLDHDLAFSRLIVELRRLHPGHV 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 LVLAVSNAFC---LGIG---FVLQ-----QNAQAHPLSDFLS-PRLLDLAMHVR---- 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 74 LPDEELQW----FVNAGWNGAMCLLHASLSEYVLLFGT-----ALGSR 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 -----RWLGGIGFVWVGVGLAVALAHGELSVEPLATNLLFALGSLSRQTKQPLGRQ 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 115 GHSGETVWHCPGEATAVWGP-----NTMVVEYGRGVIPSTLAPALADTVFSTQD 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 GWAGLALLAGGVTAFTVAGQPTGGDAVNTPLRHWV-----IIGVMVGFALLLTTTAKQS 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 165 FLTLFVTLRSYARGL 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 RLSAAPVLLSLAAGL 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-834-792-2
; Sequence 2, Application US/09834792
; Patent No. US20020037515A1
; GENERAL INFORMATION:
; APPLICANT: Mount Sinai School of Medicine of NYU
; TITLE OF INVENTION: TRP8, A TRANSIENT RECEPTOR POTENTIAL
; TITLE OF INVENTION: CHANNEL EXPRESSED IN TASTE RECEPTOR CELL
; FILE REFERENCE: AP32911 070165.0589
; CURRENT APPLICATION NUMBER: US/09/834,792
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,491
; PRIOR FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Murine TRP8
US-09-834-792-2

Query Match      8.2%; Score 83; DB 9; Length 1158;
Best Local Similarity 22.2%; Pred. No. 25;
Matches 40; Conservative 17; Mismatches 53; Indels 70; Gaps 7;

QY 29 WLGTQSF-VFQREIEIAQLARQYAGLDHDLAFSRLIVEL-----RRLH----- 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 WILTSALHVGLARHVQAVR-----DHSLASTSTKIRVVAIGMASLDRLHRLLDGVHQ 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 70 ----PGHVLPEELQWVFVNAGWNGAMCLLHASLSEYVLLFGTALGS----- 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 KEDTPIHYPADE-----GNIQGPLCLDSNLSHFILVESGALSGNDGLTELQLSL 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 114 -----RGHSGETVWHCPGEATAVWGPNT-----NMVVEYGRGVIPSTLA 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 EXHISQRTGYGTCIQIPVLCLLVNGDPNTILERSAVEQAAPWLLLAGSGGIADVLA 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-10-026-188-5
; Sequence 5, Application US/10026188
; Publication No. US20020164645A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
; TITLE OF INVENTION: Ion Channel
; FILE REFERENCE: 02307E-114910US
; CURRENT APPLICATION NUMBER: US/10/026,188
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/259,379
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse ltrpc5 predicted amino acid sequence
US-10-026-188-5

Query Match      8.2%; Score 83; DB 13; Length 1158;
Best Local Similarity 22.2%; Pred. No. 25;
Matches 40; Conservative 17; Mismatches 53; Indels 70; Gaps 7;

QY 29 WLGTQSF-VFQREIEIAQLARQYAGLDHDLAFSRLIVEL-----RRLH----- 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 WILTSALHVGLARHVQAVR-----DHSLASTSTKIRVVAIGMASLDRLHRLLDGVHQ 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 70 -----PCHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGS----- 113
Db 157 KEDTPIHYPADE-----GNIQGPLCLDSNLSHFILVESGALGSGNDGLTELQLSL 207
QY 114 -----RGHSGETVVGHPGGEATAVEWGPNT-----WMVEYGRGVIPSTILA 152
Db 208 EXHISQORTGYGTSIQIPVLCLLVNGDPNTLERISRAVEQAAPWLLILAGSGGIADVLA 267

RESULT 15
US-10-794-897-4
; Sequence 4, Application US/10794897
; Publication No. US20040259160A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Jeffrey D.
; APPLICANT: Zhou, Yun-Ping
; APPLICANT: Metabolex, Inc.
; TITLE OF INVENTION: Methods and Compositions for Treating and Diagnosing
; TITLE OF INVENTION: Diabetes and Related Diseases Involving Beta-TRP
; FILE REFERENCE: 016325-019010US
; CURRENT APPLICATION NUMBER: US/10/794,897
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: US 60/452,596
; PRIOR FILING DATE: 2003-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse betaTRP (TRPM5) calcium channel
US-10-794-897-4

Query Match 8.2%; Score 83; DB 16; Length 1158;
Best Local Similarity 22.2%; Pred. No. 25;
Matches 40; Conservative 17; Mismatches 53; Indels 70; Gaps 7;
QY 29 WLGTQSF-VFQREETIAQLARQVAGLDHDLAFSRLVEL-----RRLH----- 69
Db 102 WILTSALHVGARHVGQAVR-----DHSLASTSTKIRVVAIGMASLDRILHRQLLDGVHQ 156
QY 70 -----PCHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGS----- 113
Db 157 KEDTPIHYPADE-----GNIQGPLCLDSNLSHFILVESGALGSGNDGLTELQLSL 207
QY 114 -----RGHSGETVVGHPGGEATAVEWGPNT-----WMVEYGRGVIPSTILA 152
Db 208 EXHISQORTGYGTSIQIPVLCLLVNGDPNTLERISRAVEQAAPWLLILAGSGGIADVLA 267

Search completed: October 29, 2005, 04:50:48
Job time : 170 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 04:37:06 ; Search time 39 seconds
(without alignments)

473.683 Million cell updates/sec

Title: US-09-823-069A-2

Perfect score: 1016

Sequence: 1 MQWAVGRRWAAALLAVAA.....RSVAGRLRLTYLFGQDP 192

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990.5	97.5	223	JC5266	sigma receptor 1 -
2	903.5	88.9	223	JC5815	sigma receptor 1 -
3	226.5	22.3	256	T46871	C-8 sterol isomera
4	212.5	20.9	222	JH0488	C-8 sterol isomera
5	210.5	20.7	221	S33457	C-8 sterol isomera
6	206	20.3	219	T38129	C-8 sterol isomera
7	183.5	18.1	241	S29697	C-8 sterol isomera
8	94	9.3	393	B86189	protein T2N20.9 [
9	86	8.5	471	H70559	hypothetical prote
10	84	8.3	382	D75541	probable CynX-rela
11	83	8.2	401	YTECT0	tetracycline resis
12	83	8.2	644	T34879	probable integral
13	83	8.2	789	C83949	hypothetical prote
14	81.5	8.0	253	AF2310	serine acetyltrans
15	81.5	8.0	345	H86150	hypothetical prote
16	81.5	8.0	1198	T28678	polyketide synthas
17	81	8.0	385	E83506	probable MFS trans
18	80.5	7.9	460	T09956	NADH2 dehydrogenas
19	80	7.9	121	AC2869	conserved hypotet
20	80	7.9	121	E97645	hypothetical prote
21	80	7.9	415	AG2044	hypothetical prote
22	80	7.9	597	B55513	hoxx protein - Alc
23	79.5	7.8	178	A69450	anthranilate synth
24	79.5	7.8	491	AH0345	probable divalent
25	79.5	7.8	692	S37976	hypothetical prote
26	79	7.8	629	T36473	probable sodium/pr
27	78.5	7.7	250	A11981	serine acetyltrans
28	78.5	7.7	431	E82139	hypothetical prote
29	78.5	7.7	460	T09866	NADH2 dehydrogenas

30	78.5	7.7	469	2	G87513	beta-glucosidase [
31	78	7.7	1159	2	E70741	probable regulator
32	77.5	7.6	548	2	G70610	hypothetical prote
33	77	7.6	216	2	F69166	hypothetical prote
34	77	7.6	459	2	S10196	NADH2 dehydrogenas
35	77	7.6	527	2	AE2932	Na+/H+ antiporter
36	77	7.6	550	2	B98350	hypothetical prote
37	76.5	7.5	324	2	G69851	conserved hypotet
38	76.5	7.5	339	2	T29057	hypothetical prote
39	76.5	7.5	410	2	C84176	oxalate/formate an
40	76.5	7.5	471	2	AC3450	Na+ driven multidr
41	76	7.5	272	2	AD3004	hypothetical prote
42	76	7.5	689	2	H03336	conserved hypotet
43	75.5	7.4	262	2	G82280	vibriobactin-speci
44	75.5	7.4	507	2	S76687	hypothetical prote
45	75	7.4	275	2	S77462	hypothetical prote

ALIGNMENTS

RESULT 1

JC5266

sigma receptor 1 - human

C;Species: Homo sapiens (man)

C;Date: 25-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

C;Accession: JC5266

R;Kekuda, R.; Prasad, P.D.; Rei, Y.J.; Leibach, F.H.; Ganapathy, V.

Biochem. Biophys. Res. Commun. 229, 553-558, 1996

A;Title: Cloning and functional expression of the human type 1 sigma receptor (hsigmaR1)

A;Reference number: JC5266; MUID:97127440; PMID:8954936

A;Contents: placental cell

A;Accession: JC5266

A;Molecule type: mRNA

A;Residues: 1-223 <KEK>

A;Cross-references: UNIPROT:Q99720; GB:U75283; NID:g1906590; PIDN:ABE50402.1; PID:g1783

C;Comment: This receptor interacts with several antipsychotic drugs such as haloperidol

F;92-113/Domain: transmembrane #status predicted <TMM>

Query Match 97.5%; Score 990.5; DB 2; Length 223;

Best Local Similarity 86.1%; Pred. No. 1.5e-85;

Matches 192; Conservative 0; Mismatches 0; Indels 31; Gaps 1;

Qy	1	MQWAVGRRWAAALLAVAAVLTVVWMLGTQSFVQREELAQARQYAGLDHDLAFSR	60
Db	1	MQWAVGRRWAAALLAVAAVLTVVWMLGTQSFVQREELAQARQYAGLDHDLAFSR	60
Qy	61	LIVELRLHPGHVLPDEELQWVFVNAGWGMGAMCLLHASLSEVYLLFGTALGSRGHS---	117
Db	61	LIVELRLHPGHVLPDEELQWVFVNAGWGMGAMCLLHASLSEVYLLFGTALGSRGHSRY	120
Qy	118	-----GFTVVHGCEATAVEMGPNMTMVEYGRGVIPS	149
Db	121	WAFISDTIISGTFHQWREGTKSEVPYGPETVVHVGCEATAVEMGPNMTMVEYGRGVIPS	180
Qy	150	TLAFALADTVFSTQDFLTFLTYLRSYARGRLBELTYLFGQDP	192
Db	181	TLAFALADTVFSTQDFLTFLTYLRSYARGRLBELTYLFGQDP	223

RESULT 2

JC5815

sigma receptor 1 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004

C;Accession: JC5815

R;Seth, P.; Leibach, F.H.; Ganapathy, V.

Biochem. Biophys. Res. Commun. 241, 535-540, 1997

A;Title: Cloning and structural analysis of the cDNA and the gene encoding the murine t

A;Reference number: JC5815; MUID:98086386; PMID:9425306

A;Accession: JC5815

A;Molecule type: mRNA

A;Residues: 1-223 <SET>

A:Cross-references: UNIPROT:O55242; GB:AF030198; NID:q2809119; PIDN:AAC39951.1; PID:q2809119
C:Comment: This protein interacts with endogenous steroid hormones, progesterone and testosterone
C:Genetics: 51/1; 118/1; 149/1
A:Introns: 171; Domain: transmembrane #status predicted <TMM>

Query Match 88.9%; Score 903.5; DB 2; Length 223;
Best Local Similarity 77.0%; Pred. No. 2.2e-77;
Matches 171; Conservative 7; Mismatches 13; Indels 31; Gaps 1;

QY 1 MQWAVGRWMAALLIAVAALVTQVWMLGTQSFVQREIEIAQLARQYAGLDHSLAFSR 60
DB 1 MPWAAGRWAWITLITITIAVLQAAWMLGTQNFVFSREIEIAQLARQYAGLDHSLAFSR 60
QY 61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGRHS --- 117
DB 61 LIVELRLHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGHSGRY 120
QY 118 -----GETVVHGPGEATAVEMGNTMMVYGRGVIPS 149
DB 121 WAEISDTISGTHQWKGTTSEVYPGETVVHGPGEATALEWGNPTMMVYGRGVIPS 180
QY 150 TLAFALADTVFSTQDFTLTYLRSYARGLRLELTLYLFGQD 191
DB 181 TLFFALADTVFSTQDFTLTYLRSYARGLRLELTLYLFGQD 222

RESULT 3
T46871
N:Alternate names: EC 5.3.3.- [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C:Accession: T46871
R:Gilbert, J.; Orbach, M.J.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z24117
A:Accession: T46871
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-256 <GIL>
A:Cross-references: UNIPROT:O92254; EMBL:U59671; PIDN:AAB09470.1
A:Experimental source: strain OR23-74-1A
C:Genetics:
A:Gene: erg-1
A:Map position: V
A:Introns: 145/1
C:Keywords: intramolecular oxidoreductase; isomerase; sterol biosynthesis

Query Match 22.3%; Score 226.5; DB 2; Length 256;
Best Local Similarity 33.7%; Pred. No. 8.6e-14;
Matches 69; Conservative 33; Mismatches 62; Indels 41; Gaps 7;

QY 4 AVGRWAWAALLIAVAALVTQVWMLGTQSFVQREIEIAQLARQYAGLDHSLAFSR 59
DB 39 SGGWLUKFAILPALVAPYAVLEQL--ESFVFDTEHLHLDSKRAISRHNQDYKAIYK 96
QY 60 RLIVEL--RLRHPGHVLPDEELQWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGRHS 117
DB 97 YIVDELNDRNGVAPYVNNDEE--WVFNAGGAGWYIITHASITEVLIIFGTATGTEGT 154
QY 118 G-----ETVVHGPGEATAVEMGNTMMVYGRGVIPS 146
DB 155 GRHTADDYFHLTGTQYAVPGEPEYVPPGVSVHHLVRGTVKQVRMPESCFALEYPRGW 214
QY 147 IPSTLAFALADTVFSTQDFTLTYLRSYARGLRLELTLYLFGQD 171
DB 215 IPMLFFGYADTSLSTDFPFLWRT 239

RESULT 4
JH0488
C-8 sterol isomerase - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YM8325.03; protein YMR202w
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JH0488; S59443
R:Arthington, B.A.; Hoskins, J.A.; Skatrud, P.L.; Bard, M.
Gene 107, 173-174, 1991
A:Title: Nucleotide sequence of the gene encoding yeast C-8 sterol isomerase.
A:Reference number: JH0488; MUID:92077431; PMID:1743517
A:Accession: JH0488
A:Molecule type: DNA
A:Residues: 1-222 <ART>
A:Cross-references: UNIPROT:P32352; EMBL:M74037; NID:g171472; PIDN:AAA34593.1; PID:g171472
R:Odell, C.; Bowman, S.
submitted to the EMBL Data Library, March 1995
A:Reference number: S59441
A:Accession: S59443
A:Molecule type: DNA
A:Residues: 1-222 <ODR>
A:Cross-references: EMBL:Z48755; NID:g736296; PIDN:CAA88643.1; PID:g736299; MIPS:YMR202w
A:Experimental source: strain AB972
C:Comment: This enzyme catalyzes the reaction which results in unsaturation at C-7 in the
C:Genetics:
A:Gene: SGD:ERG2
A:Cross-references: SGD:S0004815; MIPS:YMR202w
A:Map position: 13R
C:Keywords: transmembrane protein
F16-12/Domain: transmembrane #status predicted <TMM>

Query Match 20.9%; Score 212.5; DB 2; Length 222;
Best Local Similarity 30.0%; Pred. No. 1.5e-12;
Matches 62; Conservative 36; Mismatches 66; Indels 43; Gaps 8;
QY 14 LLLAAVAALVTQVW--LWLGTSFVQREIEIAQLARQYAGLDHE---LAFSLRILVELRLH 69
DB 8 LLIGVGVYIMNVLTFTLPT--NYMFPKTIANEICNSVSKHNAEGLSTEDLLQDVORDAL 66
QY 70 PGHVLPEEL-----QWVFNAGWGMGAMCLLHASLSEYVLLFGTALSGRHSGE --- 119
DB 67 ASH-YGDEYINRYKVEWVFNAGWGMGAMCLLHASVSEYLLIFGTAVGTEGTGVHPAD 125
QY 120 ---TVVHG-----PGEATAVEMGNTMMVYGRGVIPSTLTA 152
DB 126 DVFTILHGTQIAALPYATEAEVYTPGMTHLKKGYAKQVSMPPGGSFALELAQCWIPCMPL 185
QY 153 FALADTVFSTQDFTLTYLRSYARGLR 179
DB 186 FGFLDTFSSTLDTLYTETVYLTARDM 212

RESULT 5
S33457
C-8 sterol isomerase - rice blast fungus
C:Species: Magnaporthe grisea (rice blast fungus)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S44061; S33457
R:Keon, J.P.R.; James, C.S.; Court, S.; Baden-Daintree, C.; Bailey, A.M.; Burden, R.S.;
Curr. Genet. 25, 531-537, 1994
A:Title: Isolation of the ERG2 gene, encoding sterol Delta(8) -> Delta(7) isomerase, fr
A:Reference number: S44061; MUID:94363780; PMID:8082205
A:Accession: S44061
A:Molecule type: DNA
A:Residues: 1-221 <KE2>
A:Cross-references: UNIPROT:P33281; EMBL:Z22775; NID:g311321; PID:g311322
C:Genetics:
A:Introns: 101/1

Query Match 20.7%; Score 210.5; DB 2; Length 221;
Best Local Similarity 30.2%; Pred. No. 2.3e-12;
Matches 60; Conservative 31; Mismatches 65; Indels 43; Gaps 5;
QY 11 WAALLAAVAALVTQVWMLWLTQTS----FVFOREIEIAQLARQ---YAGLDHSLAFSR 63
DB 11 FVAVLLAVVSP-----VYLAERLESFYVFDHKLHSLSTQIAIAQHGNNTAIVGHIVD 65

```

Qy   64 ELRLHPGHVLPDEELQWVPVNAGGMWGAMCLLHLSSEYVLFFGTALSGRH----- 116
      ||||| : ||| ||||| : |:|||::|:|||::|:|
Db   66 ELRARPETTYISVOQEWEVFNNAGGAMGGMYIIHASVTLYLIIFGTAIGTEGHTGRHTAD 125
      ||||| : ||| ||||| : |:|||::|:|||::|:|
Qy   117 -----SGETVVHGPGE-----ATAVEWGPNNTWMVSYVRGVIPSTILA 152
      :||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db   126 DYNFNLITGEQWAYVFGYEYEVYPAGSVHHLRGRGTKVKQKMPGCCFALEVARGWIIPPMLP 185
      :||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy   153 FALAADVTFSTDFTLTYYT 171
      ||| : ||| ||||| : |:|||::|:|||::|:|
Db   186 FGPDGLGSSTLDPTTLMQT 204
      ||| : ||| ||||| : |:|||::|:|||::|:|


RESULT 6
T38129
C-8 sterol isomerase - fission yeast (Schizosaccharomyces pombe)
C-Species: Schizosaccharomyces pombe
C&Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C&Accession: T38129
R&Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1997
A&Reference number: Z21772
A&Accession: T38129
A&Status: preliminary; translated from GB/EMBL/DDDBJ
A&Molecule type: DNA
A&Residues: 1-219 <BAD>
A&Cross-references: UNIPROT:P87113; EMBL:Z95334; PIDN:CABO8601.1; GSPDB:GN00066; SPDB:SFE
A&Experimental source: strain 972n-; cosmid c20G8
C&Genetics:
A&Gene: SPDB:SPAC20GB.07c
A&Map position: 1
```

	Query Match	20.3%;	Score 206;	DB 2;	Length 219;	
	Best Local Similarity	32.9%;	Pred. No. 6e-12;			
	Matches	55;	Conservative 22;	Mismatches 52;	Indels 38;	Gaps 4;
Qy	39	REETAQLARQYAGLDHSLA	SRLIVELRRLLHPGHVLPD	ELOWFVNAGGAMGCILHA	98	
Dd	42	KQSIALYANTKALLYDL	S---DRLVAEYGD---	ITPVNDEWHNNAGGMTILHA	96	
Qy	99	SLSEYVLFLFGTALSGRHS	G-----	ETVVHGGEATAVEWGP	135	
Dd	97	SFSLEYLFPGTPICTEGSH	GVHMADDFYTILRGQLAAS	NDL EARVYLPFGDHVPWG	156	
Qy	136	NT-----	NNVYEGRGVIPSTLAPALAD	TVFSTQDFLTLYPTL	172	
bh	157	TAOYSMPSGBCFALELAOG	WIYVSMLPFGFMWDGLF	FSTIDFGTYKYTV	203	

RESULT 7
S29697
C-8 sterol isomerase - smut fungus (Ustilago maydis)
C:Species: Ustilago maydis (corn smut)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S29697
R:Burden, R.S.; James, C.S.; Bailey, A.M.; Keon, J.P.; Croxon, R.; Bard, M.; Hargreaves, S.
submitted to the EMBL Data Library, October 1992
A:Description: Isolation and characterization of the ERG2 gene, encoding delta 8 - delta
A:Reference number: S29697
A:Accession: S29697
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <BUR>
A:Cross-references: UNIPROT:P32360; EMBL:Z17311; NID:g2966; PID:g2967
C:Genetics:
C:Introns: 142/3

Query Match 18.1%; Score 183.5; DB 2; Length 241;
Best Local Similarity 25.9%; Pred. No. 8.7e-10;
Matches 58; Conservative 35; Mismatches 64; Indels 67; Gaps 8;
Ov 7 RRW-----AWAALLLAAVAVLTQVVMWLGTQSFVFORBETAQLA-----ROYAG--- 51

Db 21 RSMIIVSAALVGFCAIALDLSIRSF-----YIFDHKAIYKIATAVANHPGNAT 71
 QY 52 -----LDHELAFSRLIVELRLHPGHVLPDBELQVFNAGGMMGAMCLLHASLSEYVL 105
 Db 72 AIFDDVDLNRADPKLAPYKNH-----FSDSEWMFNAGGAMGSMFIHASVTEYLI 126
 QY 106 LFGTALGSRGHGE-----TVVHG-----PGETAATVWGP----- 135
 Db 127 FFGTGVGTGHTGRHTADDYFNILTCNQYAFAPAGALKAEHYFAGSVVHLLRRGTVKQYMP 186
 QY 136 --NTWNVYGRGVIPSTLAFALADTVFSTQDFLTFLFYTLRSYAR 177
 Db 187 EDGCWALELAQGWIPPLPFGGLADVLSSLDLPFTFGITVWITAR 230

RESULT 8
 B86189
 protein T25N20.9 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B86189
 R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizlar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86189
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <STO>
 A:Cross-references: UNIPROT:Q9ZVZ4; GB:AE005172; MID:g8778729; PIDN:AAF979737.1; GSPDB:B
 C:Genetics:
 A:Gene: T25N20.9
 A:Map position: 1

Query Match 9.3%; Score 94; DB 2; Length 393;
 Best Local Similarity 24.9%; Pred. No. 0.38;
 Matches 51; Conservative 28; Mismatches 54; Indels 72; Gaps 11;

QY 14 LLLAVALVLTQVVMWLGTO---SFVFOR-----EEIAQLA--ROYAGLDHELAFSRL 61
 Db 160 LLIVWFUKVTLVTLGLVGLGFSWVLEEVKPEFTTEIVRNAYERTQADLD-----L 212
 QY 62 IVELRLRHPGHVLPDE-----BLQWVFNAGGMMGAMCLLHASLSEYVL 105
 Db 213 GVKLK-----LLEDFKGFVTSRKFSRCTGSDSKWKINQDGPMLNSKCVLYKSAIEVS 266
 QY 106 LFGTALGSRG--HSG-----ETVHG-----PGETAATVWGP----- 134
 Db 267 IWCWPLQTSGLFHTGFFSSSITVLGRVTEWTEGKFCGYSIHETNWSGKTKWSTVLQLD 326
 QY 135 PNTWNVYGRGVIPSTLAFALADTV 159
 Db 327 PNTWVLEYSSQSSVNDSSLLSLTI 351

RESULT 9
 H70559
 hypothetical protein Rv1634 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: H70559
 R:Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998

A:Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70559
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-471 <COL>
A:Cross-references: UNIPROT:006151; GB:Z95554; GB:AL123456; NID:g3261771; PIDN:CAB08887.
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1634
C:Superfamily: tetracycline resistance protein

Query Match 8.5%; Score 86; DB 2; Length 471;
Best Local Similarity 25.3%; Pred. No. 2.6;
Matches 41; Conservative 18; Mismatches 55; Indels 48; Gaps 8;

QY 9 WAWA---ALLAAVAALTVQV-----RRLLH-----PGRVLPDEELQWVFVNAGGWMG 91
DB 172 WNAFGVWTLTALWMLVFPVALGAGVGPGGETPVGSTHKVPVWSSLMLMGAAALISVA 231

QY 41 EIAQLARQYAGLDHAFSLRLLVEL-----RRLH-----PGRVLPDEELQWVFVNAGGWMG 91
DB 232 ALPNVILVQTAGL---LAAAALLVAVFVVVDWRIHAVALPPSVFGSGPLKWIYLTW----- 283

QY 92 AMCLLHASLSEYVLLFGTALSGRHSGETTVVHGPGGATAVEW 133
DB 284 SVQMIAAMVDTVVPLFGQRL---GHLLTPVAGFLGAALAVGW 322

RESULT 10
D75541
probable CynX-related transport protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75541
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75541
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <WHI>
A:Cross-references: UNIPROT:09RX9P; GB:AE001887; GB:AE000513; NID:g6457928; PIDN:AAF0984
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0260
A:Map position: 1
C:Superfamily: cynX protein

Query Match 8.3%; Score 84; DB 2; Length 382;
Best Local Similarity 32.9%; Pred. No. 3.2;
Matches 49; Conservative 10; Mismatches 38; Indels 52; Gaps 9;

QY 9 WAW---AALLAAVAALTVQVVMMLMGTSQSFVQREIEAQ 44
DB 167 WAWLSGAALLAWPAMLGPARSAGPAVRGSPVWSNPATLPVTLYMGTSQSLVF-FTWLTW 225

QY 45 LAR-QYAGLDHAFSLRLL-----IVEL-----RRLHGHVLPDEELQWVF-----V 84
DB 226 LARLLQDRGLSAAAGALLSLGNLVLQPLFTLSVPVLASRL--GNVRP-----LVFGLVAC 278

QY 85 NAGGWGAMCLLHASLSEYVLLFGTALGS 113
DB 279 NAAGLLGLTLWPTASPLPWVLLLVGAGS 307

RESULT 11
YIECT0
tetracycline resistance protein - Escherichia coli transposon Tn10

C:Species: Escherichia coli
C:Date: 13-Jun-1983 #sequence_revision 20-Sep-1984 #text_change 09-Jul-2004
C:Accession: A91505; A93481; A03507
R:Nguyen, T.T.; Postle, K.; Berstrand, K.P.
Gene 25, 83-92, 1983
A:Title: Sequence homology between the tetracycline-resistance determinants of Tn10 and
A:Reference number: A91505; MUID:84109550; PMID:6319234
A:Accession: A91505
A:Molecule type: DNA
A:Residues: 1-401 <NGU>
A:Cross-references: UNIPROT:P02980
R:Hillen, W.; Schollmeier, K.
Nucleic Acids Res. 11, 525-539, 1983
A:Title: Nucleotide sequence of the Tn10 encoded tetracycline resistance gene.
A:Reference number: A93481; MUID:83143319; PMID:6298728
A:Accession: A93481
A:Molecule type: DNA
A:Residues: 1-280 'E', 282-300, 'D', 302-329, 'E', 331-353, 'T', 355-401 <HLL>
A:Cross-references: GB:V00611; NID:g43700; PIDN:CAA23880.1; PID:g43701
C:Genetics:
A:Gene: tet
C:Superfamily: tetracycline resistance protein
C:Keywords: antibiotic resistance; transmembrane protein

Query Match 8.2%; Score 83; DB 1; Length 401;
Best Local Similarity 29.2%; Pred. No. 4.2;
Matches 38; Conservative 18; Mismatches 42; Indels 32; Gaps 7;

QY 14 LLLAAVAALTVQVVMMLGTSQSFVQREIEAQARQYAGLDH-ELAFSRLI--VELRRLHP 70
DB 45 VLLALYALMQVIFAPWLGKMSDRFGRPRVLLLSLIGASLDYLLAFSSALWMLYLGRLLS 104

QY 71 G-----HVLPD-----EELQWVFVNAGGWMGAMCLLHASLSEYVLLFGTALSGRG 115
DB 105 GITCATGAASVADTTTSASQKVWF-----GMLGASFGLG-----LIAGPIIG--G 150

QY 116 HSGETVTVHGP 125
DB 151 PAGEISPHSP 160

RESULT 12
T34879
probable integral membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34879
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z21561
A:Accession: T34879
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-644 <OLI>
A:Cross-references: UNIPROT:O69917; EMBL:AL023861; PIDN:CAA19596.1; GSPDB:GN00070; SCORI
C:Genetics:
A:Gene: SCORDB:SC3C8.04c

Query Match 8.2%; Score 83; DB 2; Length 644;
Best Local Similarity 26.9%; Pred. No. 7.3;
Matches 46; Conservative 18; Mismatches 63; Indels 44; Gaps 9;

QY 12 AALLAAVAALTVQVVMMLGTSQSFVQREIEAQAL-ARQYAGLDH-ELAFSRLIIVELRRLH 69
DB 110 AAVALLVDYVTVAVQVSAGTSALV-----SLAHLVGNWGTGLDHLQVPVSLVIVL--LG 163

QY 70 PGRVLPDEELQWVFVNAGGWMGAMCLLHASLSEYVLLFGTALSGRHSGETTVVHGPGGAT 129
DB 164 YGNLFGVRE-----AGRMFALPAYLFAAAMGLVFLVAARGLRGELPHADLHAPG--- 213

QY 130 AVEWGP--NTWVVEYGRGVIPSTLAFALADTVFSTQDFLTFYTLRSVARG 178

Db 214 VPLGTPGDGWL--YG-----ASLFIVLRSFANG 240

RESULT 13

C83949

hypothetical protein spoIIIE [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: C83949

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: C83949

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-789 <STO>

A;Cross-references: UNIPROT:Q9KA95; GB:AP001515; GB:BA000004; NID:gl0174886; PIDN:BA061

A;Experimental source: strain C-125

C;Genetics:

A;Gene: spoIIIE

C;Superfamily: Bacillus subtilis DNA translocase spoIIIE

Query Match 8.2%; Score 83; DB 2; Length 789;

Best Local Similarity 27.0%; Pred. No. 9.2;

Matches 57; Conservative 25; Mismatches 69; Indels 60; Gaps 12;

QY 7 RRAWAA-----LLAVAVLTQVVMWLTQSFVFOREIAQLARQYAGDHE 55

Db 8 KKVAMQSQTLFELVGLVIAVVA-LAQ-----LGTVGSTLVRLFRFFLGWYAVLSIA 61

QY 56 LAFSLVIVLRRLHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLFGTALGSRG 115

Db 62 LLVAALYIWMVRKPP-----LW-SRRIGLY--LMLLSALLSHVGLFQLOGNEG 110

QY 116 HSGETVVHGPGEATAVEMGNTWMV-----EYGRVIPSTLAPALADTVFSTQ 163

Db 111 FSDQSVIR-----NTNLFWLDWYGEVQSHDLGGMI-GAIAVAASHFLFA-- 155

QY 164 DFUTLFTLRSYARGLRLELT-----TVLFQO 190

Db 156 DGGTFLCFLFMAGLIL-LTGHISITDLFGK 185

RESULT 14

AF2310

serine acetyltransferase [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AF2310

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AF2310

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-253 <KUR>

A;Cross-references: UNIPROT:Q8YQ03; GB:BA000019; PIDN:BAB75736.1; PID:gl7133172; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all4037

C;Superfamily: Bacillus serine acetyltransferase; serine acetyltransferase homology

Query Match 8.0%; Score 81.5; DB 2; Length 253;

Best Local Similarity 25.0%; Pred. No. 3.4;

Matches 32; Conservative 19; Mismatches 34; Indels 43; Gaps 6;

QY 22 LTQVVMWLTGTSQSFVFOR-----BEIAQLARQYAGLDHDLAFSRLIVELRR 67

Db 22 LLEVLFCYCGQLLHFRVNWLYRLGLPIPLRILSHARFLTGIE----- 67

QY 68 LHPGHVLPDEELQWVFNAGGWMGAMCLLHASLSEYVLLF-GTALGSRGHS-----G 118

Db 68 IHPGATIG---QGVFIDHG--MGVVIGETAIVGDYVALIVQVTLGGTGKSGKRHPTLG 121

QY 119 ETVVHGP 126

Db 122 ENVVVGAG 129

RESULT 15

H86150

hypothetical protein F22M8.4 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: H86150

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H86150

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-345 <STO>

A;Cross-references: UNIPROT:Q9LPD0; GB:AE005172; NID:g8570442; PIDN:AAF76469.1; GSPDB:G

C;Genetics:

A;Map position: 1

C;Superfamily: arsenical pump-driving ATPase

Query Match 8.0%; Score 81.5; DB 2; Length 345;

Best Local Similarity 24.8%; Pred. No. 4.9;

Matches 35; Conservative 16; Mismatches 57; Indels 33; Gaps 5;

Search completed: October 29, 2005, 04:48:01

Job time : 49 secs

This Page Blank (uspto)